The Role of Testing in Engineering Product Development Processes

Thesis

Link(s) to article on publisher's website:
http://dx.doi.org/doi:10.21954/ou.ro.0000efbe

UNRESTRICTED

The Role of Testing in Engineering Product Development Processes

A thesis submitted for the degree of Doctor of Philosophy

Khadija Tahera
M.Sc. Cranfield University

Department of Engineering and Innovation
The Open University

DATE OF SUBMISSION: 23rd May, 2014
DATE OF AWARD: 16 SEPTEMBER, 2014

----- To my parents -----

Declaration

This thesis is the result of my own research and does not include the outcome of collaborative work, except where stated otherwise. The dissertation has not been submitted in whole or in part for consideration for any other degree.

This document contains fewer than 80 figures and fewer than 68,000 words.

Khadija Tahera
The Open University
23rd May 2014

Acknowledgements

I wish to express my gratitude to Professor Chris Earl and Professor Claudia Eckert for supervising this research and providing all kinds of support throughout my study.

I would also like to thank all the personnel who gave their valuable time to support this research.

Finally, I would like to thank my husband, my daughter and my family for supporting me.

Publications of thesis research

Tahera, Khadija; Earl, Chris and Eckert, Claudia (2014). Integrating virtual and physical testing to accelerate the engineering product development process. International Journal of Information Technology and Management, 13(2/3) pp. 154–175.

Tahera, Khadija; Earl, Chris and Eckert, Claudia (2013). Improving overlapping between testing and design in engineering product development processes. In: ASME 2013 International Design Engineering Technical Conferences (IDETC) and Computers and Information in Engineering Conference (CIE) , 4-7 August 2013 , Portland, Oregon, USA.

Tahera, Khadija; Eckert, Claudia and Earl, Chris (2013). Optimizing overlap between testing and design in engineering product development processes. In: 23rd CIRP Design Conference, 14 - 15 March 2013, Bochum, Germany, Springer, pp. 201–210.

Tahera, Khadija; Earl, Chris. and Eckert, Claudia (2012). Integrating physical and virtual testing to improve confidence in product design. In: Workshop on Modelling and Management of Engineering Processes , 29-30 November 2012, Trinity College, University of Cambridge.

Tahera, Khadija; Earl, Chris and Eckert, Claudia (2012). The role of testing in the engineering product development process. In: Proceedings of TMCE 2012, 7-11 May 2012, Karlsruhe, Germany.

Abstract

Testing components, prototypes and products comprise essential, but time consuming and costly activities throughout the product development process particularly for complex iteratively designed products. The planning of testing is a critical challenge for these complex products for which market pressures demand shorter development times. A literature review identified that testing in the design process is a relatively under researched area. An extended case study in a diesel engine company was therefore conducted to explore how testing is integrated into the product development process and how different types of testing are planned across the stages of product development.

The first part of this research study reports the empirical study. A framework resulting from this work is proposed which identifies the entities that characterise how testing should be planned. Motivated by needs of companies and research gaps identified in the literature review, the second part of this study focuses on three key problems for planning of testing in product development process: prioritisation of testing activities, scheduling of testing activities and managing the overlapping of testing and design activities.

A method of integrating Quality Function Development (QFD) and Failure Modes and Effect Analysis (FMEA) for prioritising testing activities has been proposed, which can improve the current test prioritisation process of the company. A Multiple Domain Matrix (MDM) is created consisting of the components and associated tests of a product arranged in a format that allows the dependency and interrelationships between key parts and tests to be identified. This form of representation together with the proposed prioritisation method will improve the process of organising and scheduling the testing activities.

The final study shows how virtual testing can mediate information flows between overlapping physical tests and (re)design and mitigate the risk associated with overlapping process. The study proposes a significant modification to the existing

product development process configuration for design and testing. This reconfiguration makes explicit use of virtual testing which is an extension to Computer Aided Engineering. Virtual testing mirrors the testing process through modelling and simulation, as a distinct and significant activity. Virtual testing is used to (a) enhance and (b) replace some physical tests. Finally, this study assesses the costs and risks of overlaps and their amelioration through targeted virtual testing.

Table of Contents

List of Figures

List of acronyms

ASM	Applied Signposting Model
CAE	Computer Aided Engineering
CAM	Cambridge Advanced Modeller
CAT	Computer Aided Testing
CCR	Critical Customer Requirement
CFD	Computational Fluid Dynamics
CPI	Current Product Issues
DSM	Dependency Structure Matrix
DV	Design Verification
DVP&R	Design Verification Plan and Report
ECM	Engine Control Module
FEA	Finite Element Analysis
FM	Failure Mode
FMEA	Failure Mode and Effect Analysis
FR	Functional Requirements
GW	Gateway
HoQ	House of Quality
MDM	Multiple domain matrix
NPI	New Product Introduction
NTI	New Technology Induction
NVH	Noise Vibration and Harshness
OEM	Original Equipment Manufacturers
P&E	Performance and Emissions
PD	Product Development
PDP	Product Development Process
PV	Product Validation
QFD	Quality Function Development
RPN	Risk Priority Number
SD	Concept/System Demonstration
TO	Technical Objectives
TPT	Test Prioritising Tool
VOB	Voice of business
VOC	Voice of customer
VOR	Voice of regulation
VVT	Verification, validation and Testing

Chapter 1 Introduction

"*On April 24, 1990, the Hubble Space Telescope (HST) was launched aboard the Space Shuttle. During checkout on orbit, it was discovered that the telescope could not be properly focused because of a flaw in the optics. The HST Project Manager announced this failure on June 21, 1990. The error in the HST's mirror occurred because the optical test used in this process was not set up correctly; thus the surface was polished into the wrong shape. Later, it was discovered that tests that would have found the flaw were eliminated*" (Allen et al. 1990). To save the program, the servicing mission, already planned for 1993, became much more than a simple scheduled service call adding a further $500 million for the repair mission to an initial construction cost of $1.5 billion. Almost 25 years later, similar problems can be found in development products, with many overdue and over budget while still failing to meet their quality objectives. This thesis addresses the role of testing in the product development process.

1.1 Problem definition

The fiasco of Hubble Space Telescope (HST) illustrates a number of design and project management issues and their consequences. It highlights the importance of testing. The project had cost more than any other scientific space programme and more than nearly any other space mission at that time. Yet, it failed because key end-to-end optical testing was ignored. If the test had been performed, the project would have suffered further delay but failure could be avoided.

There can be little doubt that certainty is a good thing, up to a point (Chiles 2002). The company, Perkin-Elmer, who had been world famous for its precision optics, was responsible for the design and testing on the Hubble of the Optical Telescope Assembly (OTA), including the fabrication of the primary and secondary mirrors. The design of the telescope and the measuring instruments was performed by skilled optical scientists.

Perkin-Elmer believed that the Hubble mirrors were designed to its exact specifications, yet, a problem was found. This indicates that designing to specifications doesn't prove design is error free.

The HST project was seriously delayed. During the critical time period, there was great concern about cost and schedule. This inhibited independent tests. This is the case for almost every project that regards testing as the final piece of the development effort. As a consequence, critical tests are discarded as the project is closing and there remains no room for the design and development teams to perform any kind of rational corrective actions.

The Perkin-Elmer plan for fabricating the primary mirror, placed complete reliance on the reflective null corrector, as the only test to be used in both manufacturing and verifying the mirror's surface. NASA understood and accepted this plan (Allen et al. 1990). So, NASA management was blamed for this fragility of their design process, which eventually caused this gross error. This indicates the importance of an appropriate design process for a successful project.

Allen et al. (1990) expressed that "*the most unfortunate aspect of this HST optical system failure, however, is that the data revealing these errors were available from time to time in the fabrication process but were not recognised and fully investigated at the time. Reviews were inadequate, both internally and externally, and the engineers did not do so in sufficient detail*" (Allen et al. 1990). This emphasises the need for continued care throughout the process and consideration of independent measurements.

Lessons learnt from the HST project are vital for any complex project, as for this research. The report by Allen et al. (1990) summarised the learning from the HST project and recommended that *"there will be a period of crisis in cost or schedule during most challenging projects. The project manager must be especially careful during such periods. The project should does not become distracted and fail to give proper consideration to prudent action".*

1.2 The challenge

Things haven't changed much since 1990. The cost of testing is still an issue as a major part of the overall development cost of a product; often up to 50% is required for testing. Despite that, projects fall behind schedule and still fail to meet functionality and performance expectations (Helle & Schamai 2013). Thus, in many large and complex projects, the project manager faces the dilemma of how best to validate and verify customer and engineering requirements and often decisions are made in an intuitive manner (Shabi & Reich 2012).

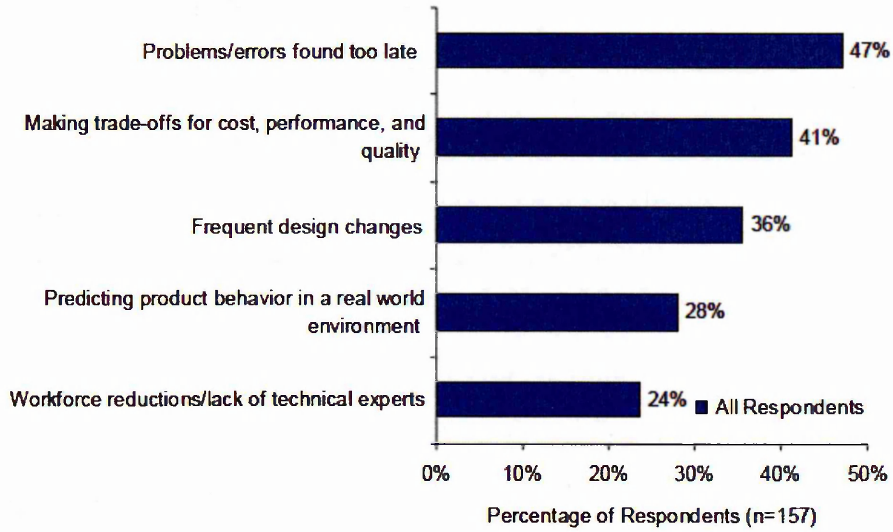


Figure 1.1 Top product design challenges (taken from (Boucher 2010))

A recent survey conducted by Aberdeen group identified top five design challenges that companies face (Boucher 2010), (see Figure 1.1). The top issue is that design problems are found too late in the development process. The costs of solving these problems increase substantially with time. The latter that failures are detected in the process the costlier they become. However, often there is no time to redesign the product because of a tight completion deadline.

To reduce the time and cost for physical testing, companies are constantly looking for alternatives such as computer aided engineering (CAE), modelling and simulation. But the replacement of physical testing through CAE is not straightforward, because these activities have different capabilities and the levels of acceptance of CAE as a means of testing rather than design analysis varies in companies. Therefore it is essential to understand the role of these CAE activities and their use in the product development.

Products are becoming increasingly complex due to rapid technological innovations, with a marked increase of electronics and software as part of what have been traditionally mechanical products. This is true for complex systems in the aerospace and automotive sectors. These complex products can no longer be treated as up-scaled versions of simple consumer products like a toasters or kettles that can be developed through a simple product development process (PDP). Complex products require intricate networks of activities in designing, testing and validating, which have complex dependencies. Proper planning of these activities is critical to achieve an effective development process. To improve the development process of complex products, it is necessary to understand different testing activities, their roles and connectivity with other activities as well as the information flow between these activities.

1.3 Motivations for this research

The principle of optimising the balance across the 'quality-cost- time' triangle has always defined design in companies but in recent years the pressure has intensified to create better products, at a cheaper price and more quickly than before. The effects on the environment and natural resources drive new technologies. National policies on environment as well as changing public values create increasing demands on products and their design. Companies, who design and produce complex and innovative products, need to consider a wide range of external drivers as well as internal business drivers to be competitive in the market. Figure 1.2 shows the top actions taken by performing businesses to address those design challenges, which were noted in Figure 1.1.

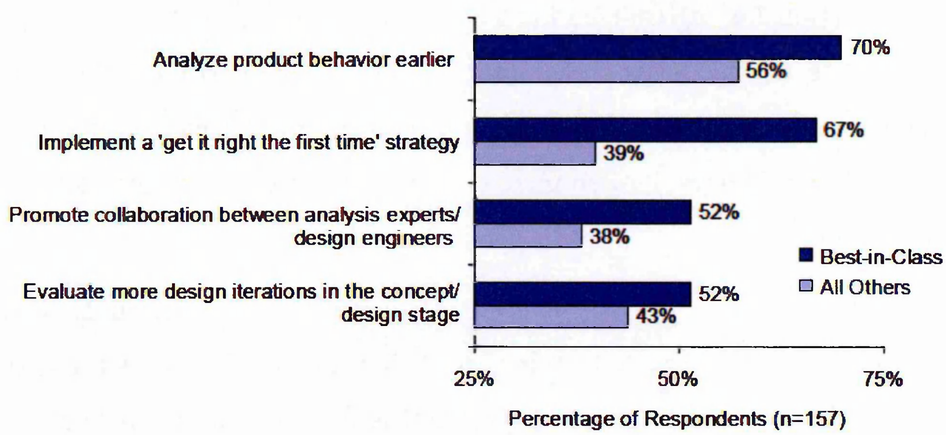


Figure 1.2 Top actions to improve the design process (taken from (Boucher 2010))

Top performing companies implement 'get it right the first time' strategy and analyse product behaviour as early as possible in the product development process. CAE approaches have been successfully used for many years for analysing product behaviour because CAE can prompt improvements before the design goes to the development stage, also CAE simulations and modelling can have a variety of other roles in supporting design and testing.

This thesis will argue that the physical testing needs to be applied as early as possible to identify design issues and to keep the relative cost of redesign to fix a discovered problem to a minimum. This means that testing should be integrated in the product development process so that each stage in the development contributes to the overall verification and validation of the product (as in Figure 1.3).

Figure 1.3 Expected process change

The integration of testing in the design process is not straightforward because the linking between product design and the testing process are often poorly understood. Companies adopt their own ways of integrating these activities in the development process, which as the case studies presented in this thesis show are often a problem. There is a clear industrial need for support with the understanding of different testing activities so that these activities can be integrated in an optimal manner.

Design analysis during the design phase is a key input to the CAE analysis and physical testing. Conversely, physical testing acts as a critical input to design and CAE, as well as providing essential validation. Understanding of these information exchanges between design and testing provides essential links in the product development process. They can identify how the testing processes can be integrated more closely with whole product development process.

1.4 Initial research questions

The work of the thesis is based on three initial questions. These questions are derived from the literature review and initial discussions with engineers and academics.

The first question was posed to understand the integration of testing in the product development process in an industrial context:

1. How is testing integrated into the product development process?

The second question investigates different types of testing and their role at different phase of the product development process

2. What are the roles of testing?

The third question explores how different testing activities are planned and scheduled through different phases of the product development process.

> 3. How are testing activities scheduled across the stages of product development process?

The aim of this research was to improve the testing process by understanding the role and characteristics of different types of testing and improving the planning of these testing activities throughout the product development process. The objective of this research is to bring new knowledge in the area of engineering testing processes that would be of benefit to industries and academia. These initial research questions are the starting point for the investigations in this thesis and inform the analysis of data from a case study undertaken as part of this PhD research. However, and perhaps more significantly, they have prompted a detailed examination of three emerging research challenges which are addressed in the latter Chapters 6-9.

1.5 Thesis structure

Figure 1.4 shows the outline of this thesis which has the following structure. Chapter 2 describes a comprehensive literature review on testing, the product development process, planning of testing and related topics. Chapter 3 explains the research methodology. Chapter 4 provides a detailed description of testing in an industry, based on an extensive case study performed at a UK based company. Building on the results of the literature review and the case study, a framework for this research in test planning is established and emerging research challenges are refined in Chapter 5. A method of prioritising testing activities is proposed in Chapter 6. A DSM based modelling approach was used to capture, visualise and analyse the complex interactions between testing and other domains, as described in Chapter 7. This is followed, in Chapter 8, by proposing an approach for minimising overlapping issues between testing and design activities. The validation of this research is explained in Chapter 9. Finally, a summary of the research contributions, limitations of this research, along with directions for further research, concludes the thesis.

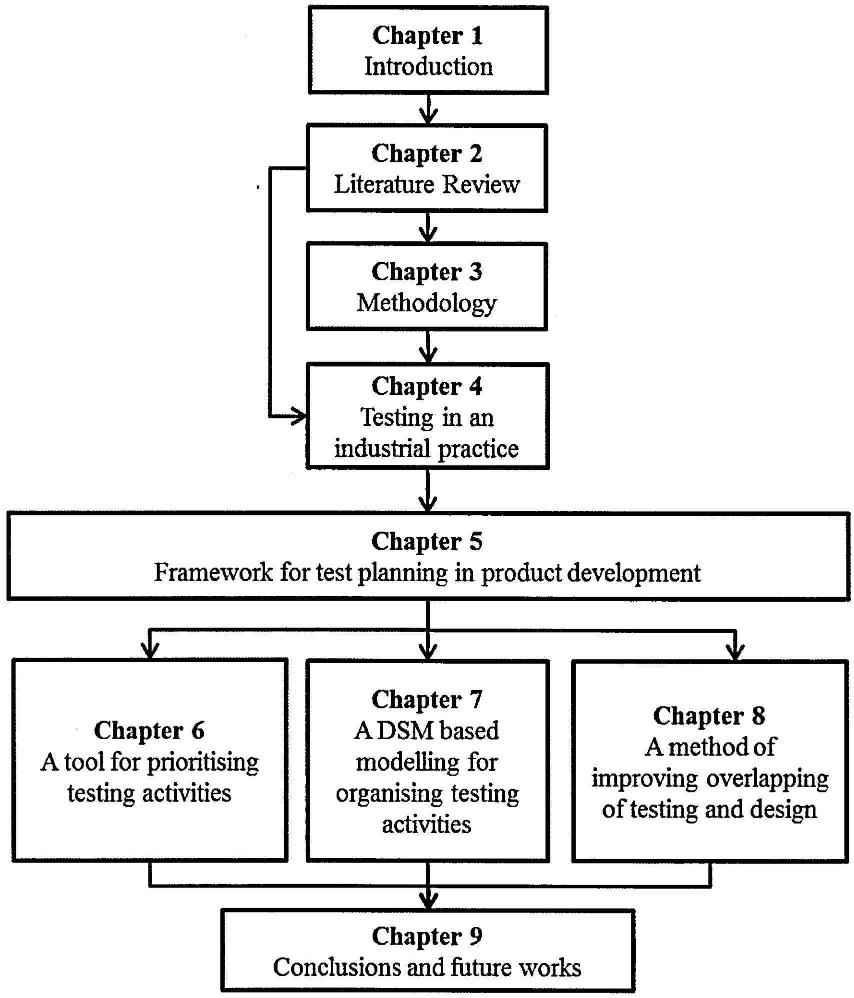


Figure 1.4 The outline of the thesis

Chapter 2 Literature Review

In design research, there has been limited attention to area of testing. Relatively a little work has been published on testing in engineering product development (O'Connor 2001, Engel 2010). Testing has received less attention than the associated design or analysis tasks of product development (Lévárdy et al. 2004).

In this chapter, first, an overview of the academic literature on testing in product development is given. Definitions of testing are then proposed in an initial attempt to answer the questions 'what is testing' along with the 'what are the types of testing' and 'what are the roles of testing'. The review of the literature on the question central to this research 'how is testing integrated with the current product development process structures' is then presented followed by consideration of the different aspects of testing planning including relevant tools.

2.1 An overview of testing in engineering design

Shabi & Reich (2012) and O'Connor (2001) both proposed that testing, although carried out in most product development projects, is seldom done in an optimal manner. In the literature, testing appears as a set of tasks to be carried out after detailed design. Generally these tasks are to be performed at the end of the product development process. Testing is not necessarily viewed as an integral part of the whole product development process from concept specification to detail and delivery. Limited analytical effort has been made to plan testing so that it takes place efficiently across the whole of product development. O'Connor (2001) stressed that the main reason for this situation appeared to be that there is no consistent set of principles, approaches and methodology for testing.

Although the importance of testing is widely recognised in research and industry practice, testing is mostly discussed in terms of specific techniques. There are many publications, which describe methods and techniques of tests, how to design a test or how to improve a test. But there is only a limited amount of research that considers testing in a wider aspect in product development. A brief summary of the main research areas in testing is presented.

Testing has received more attention in the field of software development than in hardware/product development, which is the focus of this thesis. There are several books on software testing (Mathur 2008, Ammann & Offutt 2008, Perry 1999), testing in software development (Ould & Unwin 1986), verification, validation, and testing in software engineering (Dasso & Funes 2007) and even on comparing analysis of testing techniques (Basili & Selby 1987). However, there are few books, which explicitly focus on testing in the context of the design of an engineering product. O'Connor (2001) in the book titled "Test engineering: a concise guide to cost-effective design, development, and manufacture", briefly touched on the basic principles of testing in the engineering disciplines. According to O'Connor (2001), until then there was no book on this subject and testing is seldom taught as part of the engineering curriculum. Specialist areas are taught, for example fatigue testing to mechanical engineers and digital device testing to electronics engineers. Wider aspects, particularly the multidisciplinary and systems aspects of testing remain under-represented in the curriculum. The objective of O'Connor (2001) is to provide a background of engineering test and to describe the necessary technologies and methods that will provide a foundation for plans, methods and decisions related to testing of engineered products and systems.

Pineda & Kilicay-Ergin (2010) published research on system verification, validation and testing (VVT). Their focus of their article is to provide an overview of a series of technical and management activities necessary for VVT of complex systems. They discuss VVT methodologies and tools for complex systems.

A comprehensive book was published in 2010 by Engel (2010) titled "Verification, validation, and testing of engineered systems", which also considers testing in the VVT of engineered systems. According to the author, this is the first book, which explicitly focuses on testing in engineered systems. It discusses the compendium of VVT activities and corresponding VVT methods, which are extended beyond system design and development to be implemented through the lifecycle of systems. Systems lifecycle phases include Systems Definition, Design, Implementation, Integration, Qualification, Production, Use/Maintenance and Disposal. A key driver for Engel (2010) appears to be the author's involvement in major research project called 'SysTest', commissioned by the

European Union.

SysTest, whose full title is 'Developing Methodology for Advanced Systems Testing' ran between 2002 and 2005 with a consortium of eight companies and research institutes from six European and affiliated countries. Specific outcomes of SysTest were, (a) VVT Methodology Guidelines - a collection of 41 VVT activities and 31 methods applicable to the various systems' lifecycles, (b) a VVT Process Model – a software suite that evaluates a VVT strategy in terms of its impact on cost, schedule, and risk (Hoppe et al. 2007). This project claimed that: (1) the SysTest methods and models are generic and applicable for different industries and (2) the application of SysTest can improve the process and product quality. SysTest, probably is the largest project, which researched about systems verification, validation and testing and also produced a good number of publications resulted from SysTest, e.g. Lévárdy et al. (2004), Hoppe et al. (2007), and Lévárdy & Browning (2009).

The SysTest project, as well as Engel's associated book (Engel 2010), considered the system level of products. It identified a wide range of VVT methods from different fields and developed an analysis of the cost, time and risk of using these methods in the VVT process. However, the role of testing, especially, the physical testing and the deployment of these testing activities at various stages across the PD process were not researched in detail in SysTest.

However, in this area of testing in design and product development there were two key papers published in 2001. Loch et al. (2001) and Thomke & Bell (2001) studied testing and highlighted the issues in test planning. Both papers are focused on planning and develop quantitative models for planning the testing tasks sequentially. They use the models for cost benefit analysis.

More recently a number of researchers have completed work on verification, validation and testing of engineered products. The majority of works are published in journals or conferences but are widely scattered. The following sections present a detailed review of the literature related to testing, grouped into three categories. The first describes testing and product characteristics, the second covers the definitions of testing, and the third considers the planning of testing.

2.2 Testing and product characteristics

There are a number of studies, which mention that the types and characteristics of a product i.e. hardware or software, architecture, and complexity, can affect the testing decisions. This section discusses these factors and their effects on testing.

2.2.1 Product complexity

It is common sense that a complex product, like aeroplane or a ship, will be harder to test than a simple product like a screwdriver or a toaster. Product complexity increases with three main elements (Novak & Eppinger 2001):

i) the number of product components to specify and produce,

ii) the extent of interactions to manage between these components (parts coupling),

iii) the degree of product novelty or newness.

Each additional component requires complicated testing and validation and adds to the coordination required to ensure efficient product development (Novak & Eppinger 2001). Also, complexity can arise from the relation between the product and the design processes employed in its development. This complexity can be critical with different parts of the product development requiring coordination and scheduling to make optimal use of available resources (Clarkson & Eckert 2005).

2.2.2 Product architecture

The product architecture is considered in three related ways by (Ulrich 1995):

i) the arrangement of functional elements
ii) the mapping from functional elements to physical components;
iii) the specification of the interfaces among interacting physical components.

According to Ulrich (1995) product architecture is more than just the organisation of the physical components. Rather it is focused on which functional elements should be treated in a modular way and which should be treated in an integral way.

The focus of this thesis is not the architecture of the product, but the effects it has on testing. The architecture of the product can have an effect on how the product needs to be tested (Lévárdy et al. 2004, Loch et al. 2001, Thomke & Bell 2001, Sosa et al. 2003, Baldwin & Clark 2000). Loch et al. (2001) found that the product architecture has influence on the number of tests, as well as on the time to perform a test. They also showed that a key benefit of modular product architecture lies in the reduction of testing cost (Loch et al. 2001).

Baldwin & Clark (2000) have discussed the significance of modular product architecture on testing and the impact of system level testing on product modularity. They concluded that the value of product modularity to a company is dominated by the cost of system level tests and unless the testing strategy switches from system level to modular, companies are inhibited from investing in modular product design. However, minimising system level test presents an increased risk of finding faults later in use (Jones 2007) and individually optimised subsystems may be suboptimal when they are integrated in a complete system.

2.2.3 Product innovation

A brief introduction to product innovation will now be presented which characterises testing in different types of innovation. According to Miller & Miller (2012), innovation has four main dimensions, which can be used to characterise specific product innovations:

i) market innovation (improvement in market penetration or opening of new markets),

ii) technology innovation (improvement in technology or new technology),

iii) commercialisation innovation (improvement in product development and manufacturing processes or new methods and processes)

iv) business innovation (changes in the business model or new business models).

Miller & Miller (2012) have developed a chart (Figure 2.1), which indicates four levels in each of the four dimensions ranging from existing to wholly new advanced techniques and methods.

Miller & Miller (2012) also describe four levels of product improvement (see Figure 2.1) in each of these dimensions:

- Status Quo – no improvement,
- Base / Conservative- minor improvement leading to positive change,
- Moderate - major improvement leading to positive change
- Advanced - new techniques and methods leading to a significant positive change.

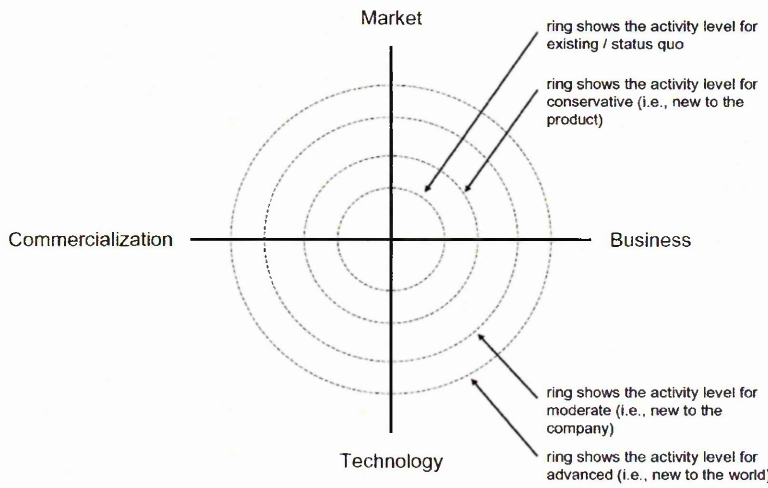


Figure 2.1 Radar chart for innovation activities in four dimensions (taken from (Miller & Miller 2012))

These four levels are often grouped into two categories when discussing product innovation, namely incremental and radical.

Incremental product innovation

Incremental product innovation perceives minor improvements or simple adjustments/ changes to existing technology. So, 'Status Quo' and 'Conservative/Base' level of improvement activities in dimensions of technology and commercialisation fall into this category.

Radical product innovation

Radical product innovation sees major improvements in technology or new techniques/ methods, which result in a new product or significant improvement in the performance of an existing product. Radical product innovation is also referred as discontinuous innovation. 'Moderate' and 'Advanced' level of improvement activities in the dimension of technology and commercialization correspond to radical innovation.

The product, technology, company and market dimensions of innovation activities tend to be responses to aspects that a company can control. On the other hand, Veryzer (1998) takes the customer's view of the product into consideration. Figure 2.2 presents product innovation as lying along two dimensions of 'technological capability' and 'product capability'. The mapping in Figure 2.2 indicates where a product innovation presents discontinuities in the customer's view.

The technological capability dimension refers to the degree to which the product involves expanding (technological) capabilities (the way product functions are performed) beyond existing boundaries. The product capability dimension refers to the benefits (functional capability) of the product as perceived and experienced by the customer or user. According to Veryzer (1998), advancement in both dimensions of product capability and technological capability can result in a radical product innovation that is both technologically and commercially discontinuous.

		Product capability	
		Same	Enhanced
Technological capability	Same	Continuous	Commercially Discontinuous
	Advanced	Technologically Discontinuous	Technologically and Commercially Discontinuous

Figure 2.2 Discontinuities presented by product innovation (Veryzer 1998)

During product development, a company will focus on the activities that are crucial for market success of that type of product. According to Song & Montoya-Weiss (1998), the objective of incremental product innovation is to capitalise on prior knowledge. Therefore, companies emphasise the post-development stages (business and market opportunity analysis and commercialisation), which are focused on their experience with existing markets and technologies. Conversely, the objective of new product innovation is to experiment and learn iteratively from the market and technology development. Thus, the development of new products should require greater emphasis on the technical development and product testing activities in order to refine technological capabilities and deliver a new product (Song & Montoya-Weiss 1998).

The activities to be performed during product innovation can be significantly different for radical and incremental innovations. For a radically innovative product, there will be a significant number of learning and experimentation types of tests during the concept development and selection phases, towards the beginning of the project. Later during product development, a company might combine verifying, validating and testing methods. If the new product is almost same as a previous version with relatively known changes, the learning from the previous product will be used to decide the testing activities on the new product. For radical innovation data from previous products is of limited application in determining testing activities.

2.3 What is testing

There are many aspects of testing. These include user testing, use scenarios, testing in manufacturing, for instance, but the focus of this thesis is on testing during product development. As mentioned earlier, testing has been mostly studied in the context of system verification and validation (V&V) and has been viewed as a subset of verification and validation (Hoppe et al. 2007). Balci (2003) states that testing is conducted to perform verification and/or validation and Secretariat (2002) states that:

> *"testing is a method of verification. The verification is executed by one or more of the following methods: test, analysis, review of design and inspection".*

Testing clearly takes place in the context of verification and validation. The next section presents definitions of these terms drawn from multiple sources. A working definition of testing is then considered for this thesis.

2.3.1 Definition of testing, verification and validation

In this section, formal definitions of testing, verification and validation are considered from several sources. These are summarised in Table 2.1 and include the IEEE standard glossary of software terminology (September 1990), the ISO 9000 Quality management system (Hoyle 2006), the telecommunications community (Engel 2010) definitions as well as the modelling and simulation community (Balci 1998).

Table 2.1 Summary of industry definitions of verification, validation and testing.

	Testing	Verification	Validation
IEEE standard glossary	An activity in which a system or component is executed under specified conditions, the results are observed or recorded, and an evaluation is made of some aspect of the system or component.	The process of evaluating a system or component to determine whether the products of a given development phase satisfy the conditions imposed at the start of that phase.	The process of evaluating a system or component during or at the end of the development process to determine whether it satisfies specified requirements.
ISO 9000 Quality management system	Not found	Verification is a process to ensure through the provision of objective evidence that specified requirements has been fulfilled.	Validation is a process to confirm that resulting product is capable of fulfilling the requirements for the specified application or intended use where known.

Telecommunication community (Engel 2010)	(1) Physical measurements taken to verify conclusions obtained from mathematical modelling and analysis (2) Physical measurements taken for the purpose of developing mathematical models	(1) Comparing an activity, a process, or a product with the corresponding requirements or specifications (2) the process of comparing two levels of an information system specification for proper correspondence (e.g., security policy model with top-level specification, top-level specification with source code or source code with object code)	(1) Test to determine whether an implemented system fulfils its requirements (2) The checking of data for correctness or for compliance with applicable standards, rules and conventions.
Modelling and simulation community (Balci 1998)	Model Testing is ascertaining whether inaccuracies or errors exist in the model. In model testing, the model is subjected to test data or test cases to determine if it functions properly. "Test failed" implies the failure of the model, not the test. A test is devised and testing is conducted to perform either validation or verification or both. Some tests are intended to judge the accuracy of model transformation from one form into another (verification).	Model Verification is substantiating that the model is transformed from one form into another, as intended, with sufficient accuracy. Model verification deals with building the model right.	Model Validation is substantiating that the model, within its domain of applicability, behaves with satisfactory accuracy consistent with the modeling and simulation objectives. Model validation deals with building the right model.

Maropoulos & Ceglarek (2010) review the standard definitions of verification and validation in the context of both digital and physical domains and provide an analysis and classification of these activities from preliminary design, to detail design in the digital domain and the physical verification and validation of products and processes, see Table 2.2.

Table 2.2 Definitions of Verification and Validation in Digital Industries (taken from Maropoulos & Ceglarek 2010)

	Verification	Validation
V&V process-	The process of evaluating software to	The process of evaluating software

es in digital design phase	determine whether the products of a given development phase satisfy the conditions imposed at the start of that phase. The process of determining that a computational model accurately represents the underlying mathematical model and its solution The process of determining that a computer model, simulation, or federation of models and simulations implementations and their associated data accurately represent the developer's conceptual description and specifications The process of determining the degree to which a modelling and simulation (M&S) system and its associated data are an accurate representation of the real world from the perspective of the intended uses of the model The process of determining that a model accurately represents the developer's conceptual description of the model and the solution to the model Providing objective evidence that the design outputs of a particular phase of the software development lifecycle meet all of the specified requirements for that phase	during or at the end of the development process to determine whether it satisfies specified requirements. The process of determining the degree to which a model is an accurate representation of the real world from the perspective of the intended uses of the model The process of determining the degree to which a model, simulation, or federation of models and simulations, and their associated data are accurate representations of the real world from the perspective of the intended use(s) The process of determining that an M&S implementation and its associated data accurately represent the developer's conceptual description and specifications The process of determining the degree to which a model is an accurate representation of the real world from the perspective of the intended uses of the model Confirmation by examination and provision of objective evidence that software specifications conform to user needs and intended uses, and that the particular requirements implemented through software can be consistently fulfilled
V&V processes in physical world	Confirmation, through the provision of objective evidence, that specified requirements have been fulfilled Provision of objective evidence that a given item fulfils specified requirements, such as confirmation that a target measurement uncertainty can be met Pertains to the examination and marking and/or issuing of a verification certificate for a measuring system Objective evidence that a process consistently produces a result or	Confirmation, through the provision of objective evidence, that the requirements for a specific intended use or application have been fulfilled where the specified requirements are adequate for an intended use Confirmation by examination and provision of evidence that the specified requirements have been fulfilled The verification process ensures that the system implementation satisfies the validated requirements

	product meeting its predetermined requirements Validation of requirements and specific assumptions is the process of ensuring that the specified requirements are sufficiently correct and complete so that the product will meet applicable airworthiness requirements	

This thesis adopts the definition of testing from the IEEE standard glossary,

> *"an activity in which a system or component is executed under specified conditions, the results are observed or recorded, and an evaluation is made of some aspect of the system or component".*

This definition is chosen partly because it matches with the definition given by the case study company (discussed in Chapter 4).

There can be many methods of verification and validation, but this research has concentrated on testing and related tasks. These included design analysis, modelling and simulation, physical tests, and analysis of test data. This research does not consider many other aspects of verification and validation including, inspections and audit, for example.

2.3.2 Types of testing

In the literature on testing, three main characteristics of a product are considered that need testing: performance, reliability and durability and there are three corresponding kinds of testing activity.

2.3.2.1 Performance

Ullman (1997) defines product performance as the measure of function and behaviour, i.e. how well the product performs what it is designed to do. Another definition from Ulrich & Eppinger (2000) that defines product performance as "how well a product implements its intended functions". Typical product performance characteristics are speed, efficiency, life, accuracy, and noise. Performance testing measures these performance characteristics. Testing also measures the performance of components and subsystems. Osteras et al. (2006) mentions that when measuring the performance of a product one must also consider characteristics such as durability. Other research by Hanks (2009) emphasises the importance of performance testing during product changes.

It highlights the necessity of conducting risk assessment and identifying the types of tests that are needed to determine whether the modified product performs as intended in actual use conditions.

2.3.2.2 Reliability

Murthy et al. (2008) state that product reliability conveys the concept of dependability, successful operation, performance and the absence of failure. There are many but similar definitions of reliability. Murthy et al. (2008) define product reliability as,

> *"the ability of a product to perform required functions, under given environmental and operational conditions and for a stated period of time".*

Wasserman (2002) defines reliability to be,

> *"the probability of a product performing its intended function over its specified period of usage, and under specified operating conditions, in a manner that meets or exceeds customer expectations".*

There are many terms in this definition, such as 'intended function', 'specified life, 'operating conditions' and 'customer expectations' and Wasserman (2002) gives detailed descriptions of these terms. Another definition by O'connor et al. (2002) states that reliability is

> *"the probability that an item will perform a required function without failure under stated conditions for a stated period of time".*

Barnard & Consulting (2008) simplifies this definition to

> *"reliability can be defined as the absence of failure in products".*

There are three basic failure mechanisms at service: infant mortality, constant (or random) failure, and wear-out. These three failures together generate the classic bathtub curve (shown in Figure 2.3). The bathtub curve is a probability curve, displayed in Figure 2.3, does not illustrate the failure rate of a single item, but shows the relative failure rate of an entire population of products over time.

At the start-up phase, there is a high risk of failure if the products are contained design flaws, improperly manufactured, improperly assembled, or stressed beyond their design limits, is referred to infant mortality. That risk comes down rapidly. There is a low but random failure during the service due to operational issues or stressed exceeding the design limit. At the end of life, the probability of failure rapidly increases, because the products wear-out.

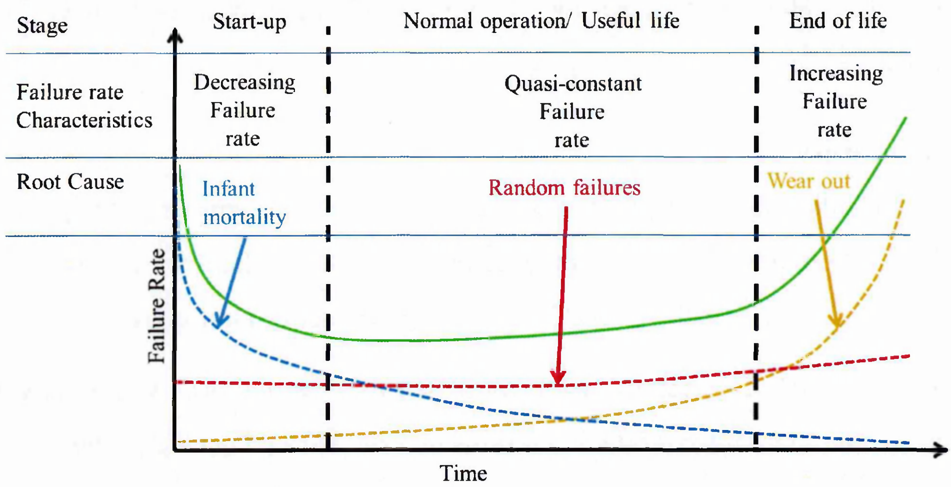


Figure 2.3 A typical Bathtub curve for hardware/product

During the development of a product, the reliability testing should focus particularly on the prevention of failure and not on the correction of failure during operation of the product. Failure prevention can be pursued by using analysis and physical testing, which can identify and eliminate both design and production deficiencies. Reliability tests include HALT (Highly Accelerated Life test), used during product design and development and HASS (Highly Accelerated Stress Screen) used during production where tests predict a product's life at high stress conditions (Nelson 2009).

2.3.2.3 Durability

Durability is usually expressed as a minimum time before the occurrence of wear-out failures. O'connor et al. (2002) states that:

> *"durability is a particular aspect of reliability, related to the ability of an item to withstand the effects of time (or of distance travelled, operating cycles, etc.) dependent mechanism such as fatigue, wear, corrosion, electrical parameter change, and so on".*

Component durability is also tested during the design phase. Conventional durability testing of engine components, for example, is carried out with full-scale prototypes in dedicated testing facilities. Testing a component can only be carried out with a small number-usually one or two-of the components. Testing to failure is needed to fully assess the durability of components, but it is rare, because of the damage that would be created on the whole engine and the test-environment during just a single failure (Holmes et al. 2011). However, approval of a final design requires a minimum number of nonfailure durability tests on the validation state of the engine design. If durability issues are identified at a late stage in the design process, redesigns are expensive and have a significant impact on delivery of the product. Therefore, Holmes et al. (2011) mention there is a need for a testing regime for jet engine components that is inexpensive, affords the possibility of obtaining a statistically significant number of data points, that can be easily performed early in the design phase of an engine, and that can facilitate reliable prediction of engine and component life.

2.3.3 Roles of testing

Alongside product verification and validation, there are several other roles of testing, which are mentioned in the research literature.

2.3.3.1 Experimentation

Thomke (2001) considered testing as a form of experimentation. According to this account, experimentation lies at the heart of every company's ability to innovate. A systematic testing of ideas can enable companies to create and refine their products. Product development project can require many experiments. Thomke (2003) defines experimentation as a process of creating knowledge which leads to the development and improvement of products, processes, systems, and organisations.

This idea of experimentation is further developed by (Erat & Kavadias 2008), and they split the experimentation process into two phases: the initial exploration phase where the design team focuses on obtaining information about the design space, and the subsequent exploitation phase where the design team, given their understanding of the design space, focuses on obtaining a 'good enough' configuration.

2.3.3.2 Uncertainty reduction

Thomke (2003) reports studies of experimentation during product development. The objective of experimentation is presented as uncertainty reduction and the objective of testing as learning:

> *"whether the product concept or proposed technical solution holds promise for addressing a new need or problem, then incorporating that information in the next round of tests so that the best product ultimately results"* (Thomke 2003).

This indicates that a key role of testing is uncertainty reduction through learning. The area of risk and uncertainty in product development process has been studied extensively in design research (for example, Yang et al. 2013, Loch et al. 2011, Wynn et al. 2011, Barrientos et al. 2010, McManus & Hastings 2006, Krishnan & Bhattacharya 2002, Tatikonda & Rosenthal 2000). A brief discussion about uncertainty is given below to provide background.

Uncertainty can be two types: Epistemic and Aleatory (Wynn et al. 2011). Epistemic uncertainty can arise from lack of knowledge or lack of definition. According to McManus & Hastings (2006), lack of knowledge refers to

> *"facts that are not known, or are known only imprecisely, that are needed to complete the system architecture in a rational way"*

Aleatory variations are sometimes termed as objective or stochastic uncertainty (Wynn et al. 2011). These are inherent variations associated with a physical system or environment - such as dimensional variation in machined components. Often the manufacturing of the prototype product and production of the final product are not completely uniform, which introduces aleatory uncertainties.

2.3.3.3 Demonstration

Yadav et al. (2006) investigates how testing can demonstrate the reliability of a product. Their paper presents a systematic approach for identifying critical elements (subsystems and components) of the system being designed and deciding the types of test to be performed to demonstrate reliability.

2.3.4 Testing in the Product Development process

This section reviews the product development process (PDP) structures to identify how the testing activities of the process are positioned in the structure. There are several product development processes that can be found in the literature and PDPs can vary widely in companies. Among those, three well-known PDPs are stage-gate, spiral and V models. In this section, the emphasis lies on highlighting of the test activities in the different models, not describing the processes in details.

A stage gate process (Figure 2.4(a)) is a structured and controlled one-way process (Cooper 1990). This process systematically follows a series of sequential steps. The number of stages differs between companies. Each step is monitored by a rigid gate review. Stage-gate processes conventionally place the testing and refinement phases at the end of the process just before starting of production (Unger & Eppinger 2011).

A spiral model (Figure 2.4(b)) repeats regular steps of concept design, system-level design, detailed design, and integration and testing (Unger & Eppinger 2011). This process is flexible; the number and span of loops can vary in companies. This model also has a structure with an integration and testing phase positioned at the end of each iteration in the spiral. The essential concept of this Spiral Model is to minimise risks by the repeated use of prototypes. With each loop of the spiral, the customer evaluates the work in progress and suggestions are made for its modification (Boehm & Hansen 1998). This model can be well used in software development but for complex engineering product, several iterations of building prototypes may raise the development cost significantly.

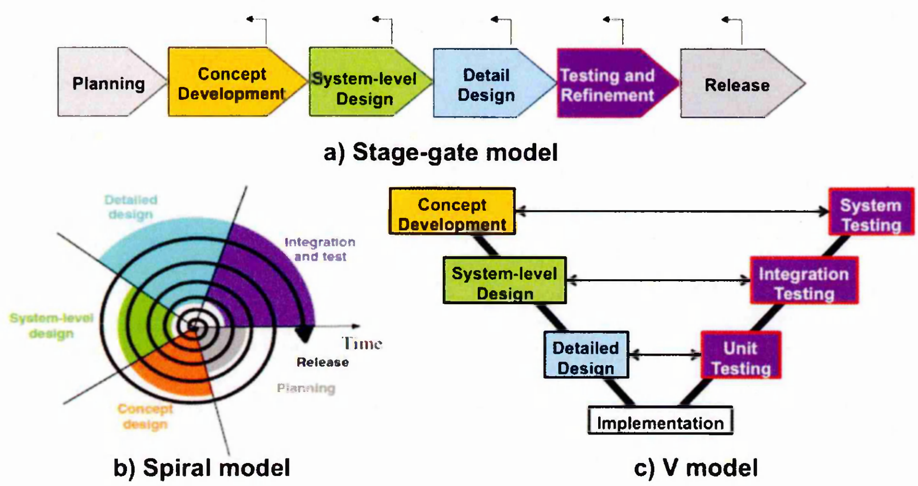


Figure 2.4 Models of product development process

The V-model (Figure 2.4(c)) is a PDP, which is widely applied in systems engineering. The “V” is a synonym for verification and validation. The V-Model demonstrates the relationships between each phase of the product and its associated phase of testing. Design and testing are ordered activities in time sequence, which complement one another across the ‘V’. So, for example, the system test is carried out on the basis of the results of concept development phase. However, this model also places the testing towards the end of the process by elaborating design activities in the left-hand side of the model and testing to the right-hand side.

Some authors have referred to the ‘fuzzy front end’ and the ‘execution-oriented back-end’ as two macro stages in the project-level development (Menor et al. 2002). The front end is considered as ‘fuzzy’ because at the beginning of the development effort information is repeatedly changeable, and consists of the activities involved in idea generation, and concept development or refinement. The back-end is the remaining portion of the development process, where the activities are involved more in implementing the chosen concept. These two 'ends' have qualitatively different objectives and characteristics and are still distinct and separate. For example, Menor et al. (2002) state that,

> *" as is the case in new product development (NPD), the front end (which is classically Marketing-centric) can become isolated from the back-end (which is classically Operations-centric), leading to 'over-the-wall' transfer of information and other dysfunctional organizational behavior".*

Often testing is considered as a task to conduct near the end of the development process and the information flow between the design (front end) and testing (which usually sits towards the backend) domains is often insufficient for an efficient product development process (Lévárdy et al. 2004). The value of the information exchange between these domains is important (Yassine et al. 2008, Unger & Eppinger 2011). However due to the pressure of producing a quality product with limited time and cost, many tasks cannot afford to wait until all required information input is arrived, and have to start 'in the dark', requiring close coordination with other interdependent activities for e.g. prototype testing and concept verification in design stage (Terwiesch et al. 2002).

Thomke & Fujimoto (2000) introduce the concept of "front-loading" as a problem solving strategy that seeks to improve development performance by shifting the identification and solving of design problems to earlier phases of a product development process. This strategy can reduce development time and cost. The authors suggest that the strategy can be achieved using approaches such as:

i) project-to-project knowledge transfer: increase the initial number of problems solved (or avoided) by more effective project-to-project transfer of problem and solution-specific information;

ii) rapid problem-solving: leverage advanced technologies (such as CAD and CAE) or other methods to increase the overall rate of development problems identified and solved.

2.3.5 Testing and design changes

Often a new engineering product is an incremental modification of an existing product, with known changes built into the testing plans. But changes also arise from (a) new customer requirements later in process or (b) identification of faults during detail design leading to urgent redesign (Eckert et al. 2009). A change of requirements from the customer has a big impact on the test plan. If a fault is identified through the test, although requirements might remain same, test must focus on the change. As every part of a design is connected at least to one other part, design changes can also be connected (Clarkson et al. 2004). The problem of change propagation can be critical for revising the testing plan. A change may be easier to incorporate in a virtual domain, whereas more effort would be required for the same change to a physical domain (Yassine et al. 2008). Managing engineering changes and especially their effect on testing is critical in the engineering company.

2.3.6 Software testing

There is a difference between the development of software and the development of hardware/physical products, such as cars, aeroplanes or engines. In developing physical products, the highest costs are experienced during the production of prototypes and the actual physical product. Whereas, the cost of producing multiple copies of software is almost negligible. Also, the test items of a physical product are likely to be discarded since the test items may get destroyed or otherwise rendered useless. However, test items of software are reusable. These can be modified and adapted to perform additional development and experiment.

Current design process models such as the V model and the waterfall model (see Figure 2.4) in software development have been considered by some researchers as inefficient in tackling the complex issues of testing (Bertolino 2009). As a consequence, more agile processes have been developed, such as, extreme programming (XP) (Myers et al. 2011). Extreme Programming is an iterative and incremental process that emphasises the importance of unit and acceptance testing in each phase of the software development process. Beck (1999) first developed the idea of the Extreme Programming, which "*...turns the conventional software process sideways. Rather than planning, analysing, and designing for the far- flung future, XP programmers do all of these activities— a little at a time— throughout development*" (Beck 2000). This idea of using an iterative process of design and testing in a shorter cycle by the repeated use of prototypes is valuable for hardware testing. However, the production of multiple copies of hardware for multiple tests is unlikely to be cost effective.

Another significant difference in testing of software and hardware is reliability testing. Software reliability does not show the same characteristics as hardware reliability. An item of physical hardware or a product can age and deteriorate. During a physical product testing, it is important to know the useful life of a component or product. Reliability testing of a product through physical testing is concerned with the history, application and use of the physical product, so that useful life of a product/component can be modelled. In contrast to this, software does not deteriorate.

Unlike hardware, software does not have an increasing failure rate as hardware does (see Figure 2.3) in the last phases of the software lifecycle (see Figure 2.5). In this phase, software approaches to obsolescence; therefore, there is no motivation for any upgrades or changes to the software (Pan 1999).

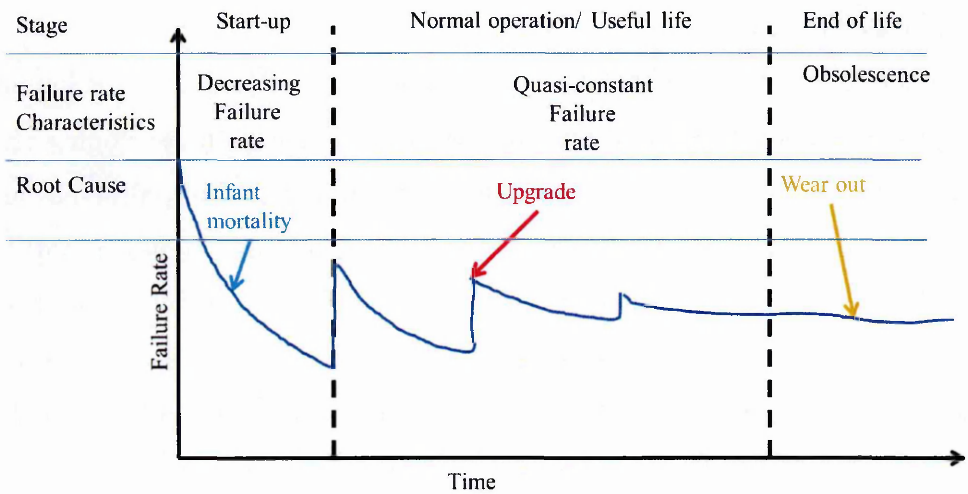


Figure 2.5 The bathtub curve for software reliability

Finally, when the design of physical product hardware is completed, the design is sent to production and product is manufactured. The manufactured product must pass a series of manufacturing tests to ensure that the manufacturing process work properly and variations in different manufactured batches of the product do not affect performance and satisfaction of functional requirements. However, software doesn't vary in batches, so, once the software is built and tested against requirements, the design of software is considered to be completed and no further testing is needed.

2.4 Planning of testing

A group of studies has considered different aspects of testing scheduling during the product development. In the early stages of product development, testing consists mainly of concept verification. Concept specification, competitive product analysis and project justifications are done in this stage. Alternative concepts are generated and evaluated. Early designs are analysed as far as is practicable to ensure the product's performance, reliability and durability against customer needs. Testing in this stage validates requirements and proves the feasibility of concept (Lévárdy et al. 2004).

The early design phases of the product development process requires special attention, as this phase can demand more coordination among design functions than other phases (Ulrich & Eppinger 2003, Ross et al. 2004). Moreover, there is always a need for a reduction in the time scale of the development cycle. This can be achieved, for example, if the quality measurement shifts towards early stages and can be implemented using upfront analysis at the concept stage (Ulrich & Eppinger 2003, Van der Auweraer & Leuridan 2005, Ayag 2005). However, the fidelity and reliability of information available in the concept stage may be limited. Testing, at concept stage, identifies the aspects of the design with high risk and uncertainty as areas for future investigation, particularly physical testing.

Where a new product is the development of an existing product, previous knowledge and analysis can be used (Ulrich & Eppinger 2003). Starting from a validated version of an existing product can help minimise product development cycle time. Timescales can be further reduced by shifting testing activities, and thus quality measurement, towards conceptual stages (Van der Auweraer & Leuridan 2005, Ayag 2005, Austin et al. 2001). One way to implement this is through modelling in a 'virtual testing' mode which also produces overall timescale reduction.

2.4.1 Sequential and parallel testing

Loch et al. (2001) have suggested that an important managerial problem in product design is the extent to which testing activities are carried out in parallel or in series. Parallel testing has the advantage of progressing more rapidly than serial testing but does not take advantage of the potential for learning between tests, thus resulting in a larger number of tests. Loch et al. (2001) have developed an analytical model to determine an optimal balance of parallel and sequential testing. They model the optimal testing strategy as a function of testing cost, prior knowledge, and testing lead-time. They apply information theory to measure the test efficiency, and show that in the case of imperfect testing (due to noise or simulated test conditions), the benefits of parallel testing decreases.

Thomke & Bell (2001) revealed that a fundamental problem in managing product development is the optimal timing, frequency, and fidelity of sequential testing activities that are carried out in order to evaluate novel product concepts and designs. They developed a mathematical model that considered testing as an activity that generates information about technical and customer-need related problems. Analysis using the model yields several findings:

i) optimal testing strategies depend on the increasing cost of redesign, the cost of a test as a function of fidelity, and the degree of correlation between sequential tests,

ii) the optimal number of tests is proportional to the square root of the ratio of avoidable cost and the cost of a single test and

iii) the relationship between sequential tests can have an impact on optimal testing strategies.

If sequential tests are increasing refinements of one another, managers should invest their budgets in a few high-fidelity tests, whereas if the tests identify problems independently of one another it may be more effective if developers carry out a higher number of lower-fidelity tests.

Erat & Kavadias (2008) have analytically modelled the implication for learning of sequential testing. The study explicitly considers the design space structure (e.g. the similarities among different design configurations) and the resulting correlations among design performances, and examines their implications for learning. The authors propose that test continuation is optimal when the previous test outcomes range between two thresholds. Test outcomes below the lower threshold indicate an overall low performing design space, where subsequent tests are unprofitable. Test outcomes above the upper threshold, on the other hand, indicate that the design team cannot expect a much higher value (than the highest value already found) by undertaking additional experiments.

In a similar way to Thomke & Bell (2001), Al Kindi & Abbas (2010) modelled testing as an activity that generates information about technical errors and customer-needs related problems that necessitate product redesign. According to Al Kindi & Abbas (2010), optimal testing strategies need to balance the trade-off between several variables, including the cost of a test, the increasing cost of redesign when discovered at a later stage, and the input/output relationship between sequential tests. They proposed that the relation between design errors and rework cost has an important role on determining the optimal testing policy. An optimal sequential testing policy is based on the condition that the net marginal value obtained from a test must be higher than the cost associated with it.

A complementary body of work sets out to analyse parallel testing. Dahan & Mendelson (2001) have quantified the costs and expected benefits of conducting parallel concept tests at the fuzzy front end of new product development. Further, their statistical model identifies that the optimal number of concepts is the ratio of the scale of profit uncertainty to the cost per concept test, i.e., higher profit uncertainty and lower testing costs increase the number of concept tests.

2.4.2 Overlapping the testing and subsequent design tasks

Sequential activities can be overlapped to reduce overall time. The overlapping of product development stages has become a common practice (Terwiesch & Loch 1999, Clark & Fujimoto 1987, Chakravarty 2001, Gerk & Qassim 2008, Roemer & Ahmadi 2004, Wang & Lin 2009, Krishnan et al. 1997). Overlapping (Figure 2.6 (b)) occurs when a downstream activity starts before an upstream activity is completed in order to reduce overall development time.

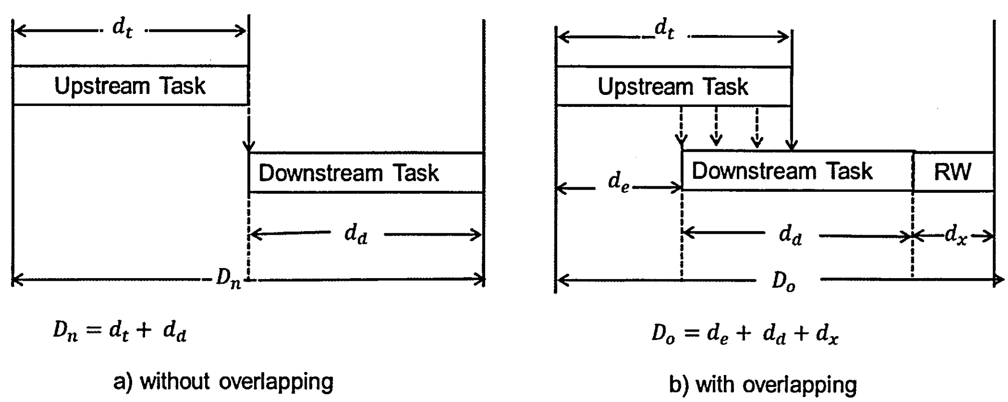


Figure 2.6 (a) tasks without overlapping, (b) with overlapping. RW indicate rework

However, researchers suggest some drawbacks of overlapping tasks during product development. In overlapping (Figure 2.6 (b)), a downstream task starts on preliminary information available from an upstream task. Rework, shown in Figure 2.6 (b) as RW, is necessary to accommodate upstream changes (Krishnan et al. 1997, Loch & Terwiesch 1998, Roemer et al. 2000, Gerk & Qassim 2008). If the uncertainty in this preliminary information is high, it can be necessary to rework most of downstream task, which reduces the benefit of overlapping (Krishnan et al. 1997, Lin et al. 2010). In the worst case, development costs may increase and product quality may worsen (Krishnan et al. 1997).

Overlapping in product development tasks and stages has received significant attention and many studies have been completed on how to optimise the overlapping process in terms of:

i) time and cost trade-offs (Chakravarty 2001, Roemer et al. 2000, Roemer & Ahmadi 2004),

ii) measuring the effectiveness of overlapping activities (Terwiesch & Loch 1999) ,

iii) conceptual framework for managing overlapping (Krishnan et al. 1997, Bogus et al. 2006),

iv) assessing risks and uncertainties in overlapping process (Yang et al. 2013, Wang & Lin 2009).

Among these studies on overlapping, only (Tahera et al. 2013, Qian et al. 2010) investigate the overlapping strategies of testing and design process. The authors claim that the testing strategies in an overlapped process differ from those in a sequential process. An analytical model for the scheduling of tests in overlapped design process was presented by Qian et al. (2010), which provided a model to find the optimal time elapsed between beginning the upstream tests and beginning the downstream development as well as the optimal downstream testing duration. They mathematically analysed that the overall cost is first convex then concave increasing with respect to upstream testing duration, and prove that there exists a unique optimum that minimises the overall cost. However, factors like problem solving capacity or opportunity cost might have impact on those conclusions. Qian et al. (2010) analytical model using opportunity costs is based on the assumption that,

> *"with increasing amount of testing, the number of residual design problems tends to decrease, and thus the rate of discovering and solving design problems should decrease".*

Although, this statement seems analytically logical but whether it is generally applicable in companies remains to be established. One issue is that accurate assessment of relevant factors and the parameters is required before developing an optimal solution.

2.4.3 Overall Verification, Validation and Testing (VVT) planning

As mentioned earlier (see section 2.3.1), testing planning has been studied as a part of the VVT planning. Outcomes from the SysTest project, Hoppe et al. (2007) and Hoppe et al. (2004) developed an approach to support strategic verification and validation planning. This approach supports the trade-off between different VVT strategies by calculating the effect of the strategy on cost, schedule, and quality.

Lévárdy et al. (2004) introduce aspects of adaptive process modelling and apply them to the VVT process to explore how increasing both project control and process flexibility can lead to improved testing process performance and higher overall project value. An adaptive process model was proposed that adjusts which activities will be done based on the state of the project and takes account of both uncertainty and ambiguity by incorporating post-VVT activity decisions. These decisions evaluate the state of the project to determine which activities will provide the maximum increase in the overall project value (Lévárdy et al. 2004).

Zakarian (2010) has developed models and algorithms for the performance analysis of product validation and test plans, especially focused on uncertainty. These models evaluate the trade-offs between the various performance criteria of the validation plan, e.g., on-time performance, cost (e.g. number of test prototypes used) and quality (number and types of requirements validated). These models are applied and validated in an industrial example through performing the performance analysis of vehicle validation and test schedule.

Recent work by Shabi & Reich (2012) has taken the approach of validating a given set of customer requirements and engineering requirements. It highlights that the literature does not offer an effective approach for associating VVT methods to VVT activities in order to satisfy customer and engineering requirements. The authors state that it is feasible to take optimal decisions about VVT for smaller projects with limited requirements, but for larger projects with a huge number of requirements, it is difficult to evaluate the overall impact of VVT methods implemented in the project. It is also necessary to evaluate the risk involved in performing or avoiding any VVT method.

Shabi & Reich (2012) propose a quantitative model for selecting an appropriate VVT approach depending on the phase of the product in the system hierarchy. The authors claim that their model is independent of project size, can structure the decision process, and therefore optimal VVT methods can be achieved in given cost and risk constraints. The use of the model has been demonstrated on a sample problem of a new payload on an aircraft.

2.4.4 Switching from physical testing to virtual testing

The time available for product development is becoming shorter. At the same time, the complexity of the product is increasing. A significant amount of the available product development time is taken up by various physical tests. Besides the time involved, these tests are generally expensive. These are two main issues: shorter product development time and cost of physical testing, which forces the consideration of alternatives to physical testing.

To reduce the increasing cost and time of testing, there is a shift in industry towards computer aided engineering (CAE) simulation. Options include computer aided testing (CAT), computational fluid dynamics (CFD) and Finite Element Analysis (FEA), for example. Advanced computing facilities have seen improvements in, speed, accuracy and economics of this computer simulation as a testing process. Advantages of using CAEs are discussed in many papers. Thomke (1998) studied the costs and benefits of computer simulation testing methods and demonstrated a 30% improvement of crashworthy testing in automotive development. A latter paper by Thomke & Fujimoto (2000) reported that the use of computer simulation tests allowed the Toyota Motor Corporation to solve about 80% of all problems by stage two (in an eight stage gateway process). As this stage was prior to the first prototypes, the CAE application resulted in about 30-40% reduction in development costs and lead time. Computer-aided design, modelling, and simulation, already allow companies to make significant advances in developing better products in less time and at a lower cost (Thomke & Reinertsen 2012).

Further, Huizinga et al. (2002) claimed that CAE tools were being used in companies to reduce the number of physical tests by simulating these tests, rather than performing them. In this paper, the authors focused on the use of CAE to simulate fatigue life related tests. The complexity of these tests varies from (relatively simple) component strength tests to full-scale vehicle fatigue life tests. Martin & de Carvalho (2006) analysed the impact of this simulation on the Product Development Process (PDP). A comparative analysis was made between PDP 'traditional model' (a term used specifically for product development without simulations or virtual tools); and the 'updated' PDP (which uses virtual simulation to analyse the project before physical prototypes are built and tested). This case study showed that simulation contributes to a reduction of time and cost of project implementation. Another study by Becker et al. (2005) identified that virtual simulation tools also have an important impact on the redefinition of the set of options to be taken into consideration in design problem-solving. Therefore, the role of physical testing is changing due to the emergence of CAE technologies (Wilkinson 2007).

Many large companies and organisations, for example, NASA is putting increasing effort into effectively using virtual simulation and CAE. NASA's Glenn Research Center initiated a collaborative effort with the aerospace industry and academia to develop its Numerical Propulsion System Simulation (NPSS), an advanced engineering environment for the analysis and design of aerospace propulsion systems and components. The study estimated that using NPSS has the potential to dramatically reduce the time, effort, and expense necessary to design and test jet engines by generating computer simulations of aerospace objects and systems. These simulations permit an engineer to "test" various design options without having to conduct costly and time-consuming real-life tests. By accelerating and streamlining the engine system design analysis and test phases, NPSS helps to bring the final product to market more quickly (Jones 2011).

2.4.5 Virtual vs Physical test in engineering design

In general, it is assumed that a physical test will provide more confidence regarding the applicability of testing data. However, there are inefficiencies in physical testing. A physical component test can deal with only limited variables and testing cannot always be comprehensive enough to include all the operating conditions (Zorriassatine et al. 2003). Furthermore physical testing is conducted in a controlled environment (Khalaf 2006a). Sometimes existing test facilities have an inappropriate testing environment and may not be the good replica of the operational environment (Hoy et al. 2008).

On the other hand, a CAE modelling can handle a spectrum of variability across many interacting variables. This can prompt the identification and correction of design deficiencies and faults. This results in a reduction in the number of prototype builds, reduction in engineering iterations, thus overall reducing cost and time. Not only analysis, but also, the advanced techniques of modelling and simulation have provided the capability of testing products virtually, which is effectively computer aided testing (CAT). These CAE analyses and especially CAT allows an earlier detection of potential engineering problems. This allows a "front-loading" on product development performance (Thomke & Fujimoto 2000).

However, the acceptance of virtual testing is still debated in industry (Tahera et al. 2012). This is because virtual tests should replicate a product's physical behaviour to provide confidence, among engineers and designers, in the virtual tests. Moreover, physical testing is a necessary industrial practice, usually required for product certification. For example, the aerospace industries need to go through a rigorous testing regime to pass certification criteria and automobile manufacturers are required to test their prototypes for regulatory and safety standards (Maropoulos & Ceglarek 2010). There are still risks in using virtual models, such as if the service conditions are not quantified. It can be difficult to correlate with the real world performance results, thus giving less confidence in the result of a virtual test (Wilkinson 2007). Moreover, some of these computer models are still poorly supported (Bringmann 2008), which makes it difficult to construct and maintain accurate and reliable models. The CAE model can contain flaws, which might be undetected until later stages in product development. In some cases CAE models can be complex in nature, so special knowledge or training is required, whilst sometimes software can be expensive to acquire thus limiting accessibility (Wasserman 2002, Chua et al. 1999).

As both virtual and physical test has advantages and limitations, the literature suggests that a combined approach of physical and virtual test might help to perform a focused test as well as increase the reliability of the test leading to a minimized iteration (Van der Auweraer & Leuridan 2005, Wilkinson 2007, Huizinga et al. 2002, Zorriassatine et al. 2003, Wasserman 2002, Chua et al. 1999, Khalaf 2006a, Khalaf 2006b). Karen et al. (2010) suggested that by feeding simulation results into physical testing results, a virtual test can be used to design an optimum system with the desired static and dynamic characteristics. This approach which mixes physical and virtual testing methods can result in a reduction in the physical prototypes built. Minimizing physical testing activities may lead to reduction in both cost and time (Wasserman 2002).

Many studies suggest correlating physical and CAE analysis data to provide an integrated solution. For example, Lee & Han (2009) have developed a Finite Element (FE) model with updating techniques to perform the correlation analysis between durability test and simulation, and implemented it into the existing FE software. A paper by Ferry et al. (2002) describes an approach where a multi-body dynamics simulation software package is trained with actual vehicle road testing data. The intention is to mirror as closely as possible the behaviour of a physical vehicle in order to assist in determining its durability characteristics under a variety of road conditions.

Wilkinson (2007) suggested that a successful project should use the merits of both virtual and physical test technologies. It was demonstrated through an example of reducing in-cab noise during the design of a new truck. Wilkinson (2007) considers that the role of physical testing is changing, is concerned with developing CAE by providing those elements that are difficult to simulate and resolving problems that slip through the product development process undetected. Also a product has many uses, and it may be necessary for it to perform in many different environmental conditions; these use scenarios can only be virtually tested because their scope and complexity would imply excessively costly and time consuming physical test.

Hence, he summarised four specific questions that should be considered before commencing physical tests are:

" 1. *Is this test necessary, or do a combination of experience and CAE prediction give sufficient confidence?*

2. *Is the test intended to simply record the output response of a system to a given set of input conditions, or is it necessary to gain an understanding of how the system delivers that response?*
3. *Is it better to use a simple CAE model to gain understanding of the system and use the physical test simply as a validation?*
4. *Does the physical test correspond to both the real-world use situation and the CAE model conditions?"* (Wilkinson 2007).

These questions highlight that there is a need to take into account the capabilities, purposes and alternatives of both physical and virtual approaches and to use both together to generate better information during testing. These critical questions provided significantly important pointers for the research presented in this thesis.

2.5 Tools used for testing planning

Companies use many tools to manage their product development processes. Among these, few assist in testing planning. Three of these tools, QFD, FMEA and DVP&R are discussed briefly.

2.5.1 Quality Function Development (QFD)

Quality function development (QFD) is a technique for translating the customer requirement to design requirements. As described by Akao (2004)

> "*QFD provides specific methods for ensuring quality throughout each stage of the product development process, starting with design. In other words, this is a method for developing a design quality aimed at satisfying the consumer and then translating the consumers' demands into design targets and major quality assurance points to be used throughout the production stage*".

Customer requirements are analysed using other tools such as the Kano model (Shahin et al. 2013) and Analytic Hierarchy Process (AHP) (Ishizaka & Labib 2011) to identify the critical customer requirements (CCRs). These requirements are then weighted by the importance of customer's wants. Accordingly, engineers identify the technical characteristics that are required to build the product as well as satisfying the CCRs. The technical characteristics are weighted against the CCRs, i.e. the degree of correspondence between a customer requirement and a technical characteristic are analysed in a House of Quality (HoQ). QFD analysis has four stages: product planning, parts deployment, process planning and process control (Hauser & Clausing 1988). QFD identifies critical technical requirements of the design which will need verification and validation by testing.

2.5.2 Failure Mode and Effect Analysis (FMEA)

FMEA was introduced in the aerospace industry in the mid-1960s, with the focus on safety issues to prevent failures and accidents. FMEA is a systematic, proactive method for failure analysis. It involves reviewing and evaluating components, assemblies and systems to identify where and how these might fail and to assess the relative impact of different failures. This analysis identifies the critical areas where design effort and resources are most needed.

A company will use FMEA to evaluate a potential design for possible failures and to prevent them by correcting proactively by changing the design rather than reacting to adverse events after failures have occurred. This emphasis on prevention may reduce risk of failure in field. FMEA is particularly useful in evaluating a New Product Introduction programme prior to implementation as well as in assessing the impact of a proposed change to an existing design. More details about FMEA and steps of FMEA analysis can be found in (Stamatis 2003). FMEA is one of the most widespread methods used in determining priorities for technical risks in the PD process especially during the testing phase (Segismundo & Miguel 2008).

2.5.3 Design Verification Plan and Report (DVP&R)

DVP&R stands for Design Verification Plan and Report. The DVP&R includes a list of verification and validation activities (can be design changes, CAE analysis and testing), starting and finishing dates of these activities, in which PD phase of the process will be performed, which item of product, what to measure and so on. DVP&R receives input from many sources, such as requirements and specifications, QFD, FMEA, robustness analysis and past history and legacy data (DFKI 2011). Costs and timing for completing tests and budget constraints all drive the DVP&R, therefore additional consideration on how and how much testing is needed (Priddle et al. 2003).

2.6 Summary

Relatively little literature has been published on testing in engineering product development, but this review shows that testing is an important topic that impacts upon most businesses. The literature has supported this view in studies across industry sectors. Testing is critical in product development of companies across a range of industries such as aerospace, automotive and electronics.

Well known product development process structures, i.e. stage-gate, spiral and V model, do not always reflect the importance of the role of testing appropriately. All of these models are limited in following ways;

i) the testing activities start very late after detailed design phase,

ii) the link between test, redesign and change tasks during the process is not clear,

iii) the interconnection between the various testing stages (in spiral model) and the types of tests used is not clear.

This literature review shows that the focus of most of the published articles on testing is about reducing costs and duration of testing. Only a few papers describe the planning and scheduling of testing tasks. Many of these include quantitative models, which consider the qualitative difference between types of tests in a limited way (except for (Thomke & Bell 2001)). Also, they have not explicitly considered at which stages of the product development process, a test is performed. Rather, a test was specified as occurring 'early' or 'late' in the process. These researches make assumptions that there is sufficient understanding of testing tasks that are used in product development. There is no paper that comprehensively describes the process of testing during product development, its planning and its different aspects in an industrial context. This thesis aims to fill that gap.

Chapter 3 Methodology

A research methodology provides a framework for systematic research. If the methodology is followed in subsequent research in a similar context, similar results will be expected. This chapter discusses the research methodology that has been followed in this thesis.

3.1 Methodological framework

In this section a methodological framework introduced by Eckert et al. (2003) is explained first, and then the steps that are followed in this research are described.

Engineering design research is inherently complex, multidimensional and multidisciplinary with threefold goals: provide an understanding of the phenomenon of design, improve the process and to produce tools and techniques for industry (Eckert et al. 2003). These goals, involve addressing four key questions (Eckert et al. 2003):

(1) How does design happen in concrete situations, which can be studied and in particular, within processes which can be improved?

(2) To what extent is it possible to understand the cognitive, social and cultural mechanisms that underlie observed the phenomena of design, by building theories?

(3) What computer tools, pencil-and-paper techniques or design methods might be useful, and how can they be developed?

(4) How can these tools or methods be introduced into industrial use, and what is their effect?

Eckert et al. (2003) propose a methodological framework for design research (Figure 3.1) which assists researchers trying to answer these four questions. This framework has eight steps.

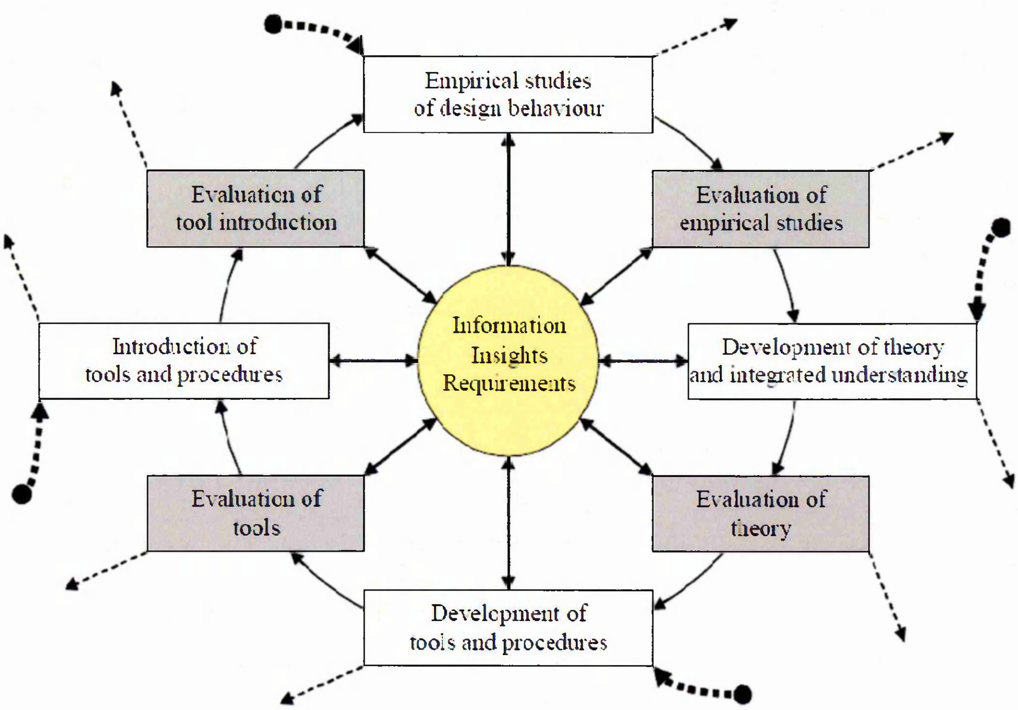


Figure 3.1 A Design Research Methodology (adopted from (Eckert et al. 2003))

An individual piece of research can enter into this cycle at any of four points (top, bottom, right, left, as indicated by heavy dotted arrows in Figure 3.1). These entry points are empirical research, theorising, tool development, and introducing tools to make changes to industrial practice. The design research methodology indicates that these activities are sequential around the cycle. However, in practice design research is often iterative, with these steps occurring in smaller iterative cycles required to verify and evaluate the understanding. Thus the steps can be performed in parallel. Essentially, this research framework is for a large group that carries on research over many years; an individual researcher cannot usually complete all these steps during the course of a PhD thesis. However, it is essential to have a clear understanding of how the thesis fits into the broader context of research in engineering design and how the findings of the thesis feed into the overarching research problems.

This PhD research enters the cycle at "Empirical studies of design behaviour" and carries on until "Developments of tools and procedures". It covers the six steps:

1. *Empirical studies of design behaviour:* this phase particularly emphasises studying the current state and identifying the criteria by which the real design is to proceed. This stage of research has provides an understanding of the various factors that influence directly or indirectly what is tested and why.

2. *Evaluation of empirical studies:* this stage validated the description of information and data in the empirical study. The evaluation process focuses on what has been leant, and how this learning can be generalised, particularly the process of learning and its fits with existing concepts of design.
3. *Development of an understanding:* the aim, at this stage, was to advance the understanding of design practice, and investigate whether this can be expressed in the form of theories or models of aspects of design or design processes.
4. *Evaluation of understanding*: these theories and models have been evaluated in terms of their methodological approaches and assumptions that have been made through reference to the empirical data in step 1.
5. *Development of methods or procedures*: new methods are developed to help the current design and testing processes and to advance the understanding of these processes
6. *Evaluation of methods and procedure*: a preliminary evaluation of proposed methods is carried out but will require development and improvement in future research.

In this research, these steps were iterative and often overlap, with empirical study, development of understanding and model building occurring simultaneously in the case study setting of this research.

3.2 Research method

This section discusses the methods which have been followed during the each phase of this research mentioned in Section 3.1.

3.2.1 Case study

To elaborate the initial understanding and to determine the factors that influence the practical environment, case studies have been carried out in this research. A case study is a qualitative research approach, which allows an investigation to retain the holistic and meaningful characteristics of the real environment (Yin 2009). Yin (2009) gives a definition, which guided the choice of using case study as a strategy for understanding testing in product development within an industrial context:

> *"a case study is an empirical inquiry that investigates a contemporary phenomenon within its real-life context; when the boundaries between*

phenomenon and context are not evident; and in which multiple sources of evidence are used" (Yin 2009).

3.2.1.1 Data collection

Interviews, emails, phone conversations and document analysis were all used to enquire about the current practice and collect empirical industry data.

Interviews

Eighteen interviews were carried out from 28th February 2011 to 21st February 2014 building on a previous series of interviews, in the same company, on system architecture reported in (Wyatt et al. 2009). Eight engineers including a senior engineer, a development engineer, a business manager, a verification and validation manager and a validation team leader (shown in Table 3.1) were interviewed.

Table 3.1 Interviews: carried out at case study company including interviewees, roles, dates and durations

Engineer	Role	Interview dates and durations(in minutes)																	
		28th Feb 2011	19th May 2011	13th Jul 2011	13th Jul 2011	23rd Aug 2011	15th Dec 2011	15th Dec 2011	3rd Feb 2012	3rd Feb 2012	27th Mar 2012	27th Mar 2012	4th May 2012	17th Sep 2012	16th Jan 2013	18th Mar 2013	9th Oct 2013	18th Oct 2013	21st Feb 2014
		127	159	76	65	135	39	148	105	40	82	75	142	119	52	117	160	151	86
Engineer 1	End to End Process Architect	X					X	X				X		X	X	X		X	X
Engineer 2	Verification and Validation Manager		X	X		X	X		X		X								
Engineer 3	Validation Team Leader - Tier 4 Interim & Final		X	X	X	X	X		X	X	X		X				X		
Engineer 4	Development Engineer			X	X														
Engineer 5	QFD and FMEA manager					X													
Engineer 6	Master FMEA owner and test planning												X						
Engineer 7	Business transformation manager													X					
Engineer 8	Core Engine mechanical system team leader																X		

The first interview, with Engineer 1, provided an overall view of testing in product development as well as an idea of the expenditure that is incurred around testing. Engineer 1 mentioned that,

> "to develop the Tier4 engines can cost R&D alone an excess of over X million, I would break it down to design and engineering is probably 15%, material is probably around 30%, and actually testing around performance is rest- around 55%. So most of the money in R&D is goes into testing for performance and durability".

The verification & validation manager (Engineer 2) and the validation team leader (Engineer 3) were interviewed to investigate how testing really happens on component, subsystem and system levels. Most of the interviews were involved with the verification and validation team who are responsible for product validation and testing. To investigate the relation between the verification & validation phases and the design phase, the meetings included development engineers in the presence of the validation engineers.

Staffs from other departments, such as FMEA, QFD, and CAE, were involved in the interviews when required. The durations of these interviews varied from forty minutes to two and a half hours. All interviews took place at the company's site. The interviews were carried out by the author. Several took place jointly with supervisors Dr Claudia Eckert and Dr Chris Earl. During the interviews, special attention was given to impartiality so that perspective of the interviewer does not limit or influence the data obtained from the interviewee. Audio recordings of these interviews were made as well as notes being taken.

The interview format was semi-structured, which is appropriate for exploratory interviews. There were no lists of fixed questions for the interviews although areas of interest and enquiry were set down beforehand. Rather, the author, as interviewer had a general list of topics to guide the interview. The initial list of topics was derived from the queries that arose from the literature review. Each interview was analysed before conducting the next interview. The analysis of each interview raised several questions, which drove the subsequent interviews. Consequently, the list of topics was refined and added to throughout the course of the interviews. Often, interviews were set up to bring in a range of engineers to gain a wider understanding and to identify gaps and shortcomings in the current process. The interviews were mostly conducted in a focused group with individuals who have specialist understanding of the testing domain (i.e. Engineer 1, Engineer 2 and Engineer 3).

The understanding and notes were generally shared and compared among the interviewees after each meeting in the company. Each of the interviews was recorded and transcribed. Some interviews were transcribed in full, whereas, the key parts were transcribed in others. However, the author has listened to each of these interviews several times to make sure that any important data were not overlooked. Once, the common ideas were observed between interviews and no new information was generated through interviews, that point was considered as saturation and deemed to be enough for data collection.

The first five interviews (from 28th February 2011 to 23rd August 2011) were discussions about a range of topics in product development, customer requirements, company process and testing, mostly in the form of fact finding. These helped to establish an understanding of the current practice in the company.

The next set of interviews (i.e. eight interviews from 15th December 2011 to 17th September 2012) focused on understanding the company's testing process through the analysis of the product development processes and identifying associated challenges that the company faced. In this stage of the investigation, specific research problems started to emerge and explicit questions were formulated according to these problems. The same group of engineers was interviewed to uncover the problems and to capture an indication of how they might be solved.

The final set of interviews (i.e. 16th January 2013 to 21st Feb 2014) mainly focused on analysis and evaluation of the proposed solutions. However, the discussions frequently shifted between problem domains and solution domains. In this way there was a recurrent and iterative refining of the understanding the current practice, associated problems and implication of the solutions. From these interviews the author gradually developed the proposed models and the suggestions for possible improvements.

Document analysis

During the meetings and interviews, software tools, diagrams, and graphic representations were presented and all provided useful information on how the company plans testing during its engine development programmes. Many documents were shared through emails. Occasionally, engineers used flip-charts to explain their conceptual frameworks, and the author took exact copies of these diagrams. The information from these diagrams was compared against that from the discussion in the interviews.

3.2.1.2 Case study method validation

This section describes the validation process of the case study research method used in this research. The validation process, suggested by Yin (2009), has been followed for validating the case study. Yin (2009) provides that the quality of a case study research design can be evaluated using four dimensions: construct validity, internal validity, external validity, and reliability. *Construct validity* can be achieved by establishing correct operational measures for the studied concepts. Using multiple sources of evidence

- interviews, observations and documents, for example, can establish the confidence in construct validity. *Internal validity* is mainly seeking to establish dependencies or relationship between conditions, that is, certain conditions are believed to lead to other conditions. This type of validity is a concern for explanatory case study and inapplicable to descriptive or exploratory case studies. *External validity* is concerned with the problem of whether a study's findings are possible to generalise to beyond the immediate case study. Maxwell (1992) explains generalizability into two contexts: internally and externally. Generalising the conclusions within the setting studied is referred as internal generalizability and beyond the setting is stated as external generalizability (Maxwell 1992). *Reliability* is ensuring that the operation of the study - such as the data collection procedure- can be repeated, with the same results. The following paragraphs describe the validity of the case study as a research method for this study.

These following four dimensions are followed in this study Table 3.2 summaries the tactics for case study validation and corresponding evidences were considered.

Table 3.2 Tactics that applied for this Case Study research validation (adopted from (Yin 2009))

Validity tests	Case Study Tactic	Phase of research in which tactic occurs	Evidence of tactic has been used
Construct validity	Use multiple sources of evidence	Data collection	Interviews and documents
	Establish chain of evidence	Data collection	Sufficient citation to the relevant portion of case study data (engineers comments), upon inspection the evidence can be revealed from data.
	Have key informants review draft case study report	Composition	Each interview started with the clarification discussion of the questions raised from previous interview. Clarification and understanding of concepts were also attained through email conversations.
Internal validity (inapplicable to descriptive or exploratory case studies)	Do pattern matching	Data analysis	Company's process models were simplified and replicated to increase the understanding. These models were used as the reference point of discussion to minimise the mental gaps, which also ensured that every one were talking about same thing.
	Do explanation building	Data analysis	N/A
	Address rival explanation	Data analysis	N/A
	Use logic models	Data analysis	N/A
External validity	Use theory in single-case studies	Research design	Intertwined nature of testing and design in product development process was seen in another companies
Reliability	Use case study protocol	Data collection	"Interview-analysis-interview-analysis" were used as a protocol
	Develop case study database	Data collection	Audio recording and notes of the interviews are stored and managed.

Construct validity has been ensured using multiple source of evidence, by establishing chain of evidence and by having the informants review on understanding. Interviews, documents were combined to collect data to understand the current practice of testing process in the engineering companies. Data source triangulation—using evidence from interviews and documents, provided verification and validity while complementing similar data. Interviewee's comments were cited throughout the thesis to establish necessary and sufficient evidence of the constructs. Each of the interviews were analysed and were discussed with the interviewees to clarify any confusion before conducting the next interview. Models of the current product development processes were created to describe the understanding of the testing phenomena occurring in the context.

As described before, an exploratory case study might not require to demonstrate the *internal validity*. However, internal validity was achieved by pattern matching. Some of the assumptions (that is, testing should be integrated in the product development process and is not a phase to be performed towards the end of the process), were made before starting the case study and were matched with the empirical findings. And Yin (2009) suggested that, this would be enough to achieve internal validity for exploratory and descriptive case study.

Similar issues, which were proposed and attempted to solve in this research, were found in other companies. This meant that the issues were not related to a standalone case in a particular company, which confirmed the *external validity*. The empirical findings from the initial case study led to theories - the testing process is closely intertwined with design in the product development, for example. Through a pair of interviews in a different company and in informal discussions with individuals in an academic institution and in a company the *external validity* was confirmed.

The *reliability* in the research work can be ensured by following the protocol of this research. The data collection protocol is described in previous section. The case study database (i.e. audio recording, notes and documents) was maintained and managed so that evidence can be established, if sought.

3.2.2 Analysis of empirical studies

Qualitative data analysis is exploratory and is useful when the phenomenon of an area needs to be understood (Creswell 2009). The qualitative analysis of empirical data provided an understanding of current practice and the underlying theories in the case study company. The analysis of interviews identified the driving factors, different dimensions and underlying reasons for the current process.

3.2.3 Development of models

The understanding from the empirical study led to a model of the system under study. Initially, the current state of the art was captured through simple models and conceptual frameworks. The company's testing process was modelled to recognise exactly where testing occurs in the product development process and provided better scope of analysis and where improvement might be made. These models were refined and developed through further discussions with engineers.

To understand the dependency and interrelationship between design and testing activities Dependency Structure Matric (DSM) was used. Dependency Structure Matric (DSM) is an established method for capturing the sequence and complex interaction of design tasks (Browning 2001). DSMs were used to capture and visualise the interdependency between components and tests.

The different testing activities that the company performs during the stages of their product development process were analysed and modelled. The testing processes were modelled using the Applied Signposting Model (ASM), which is a process modelling approach to model and simulate engineering design processes that can include iteration (Clarkson & Hamilton 2000). This process model was used to model the workflow of the testing and design activities. The Cambridge Advanced Modeller (CAM) was used as a tool for modelling and analysing the DSMs and ASMs models developed in this research (Wynn, 2011). The CAM tool was used for viewing and identifying essential relationships and to see how changes in tests may affect critical links and overall process.

3.2.4 Development of methods

A method to integrate virtual and physical testing is proposed in this research to improve the overall testing process in the company. This is a conceptual model based on the analysis of the empirical study. This method was discussed with different engineers, both in the case study company and in a different company, along with an academic expert in a university. They provided their subjective judgment, identifying the potential use of this method.

This research was an iterative process, in which an initial focus was on understanding design behaviour and defining the problem through empirical studies, followed by analysis, model building and evaluation. These later steps were largely iterative in nature, as insights from a step drove new research in the others as shown in Figure 3.2.

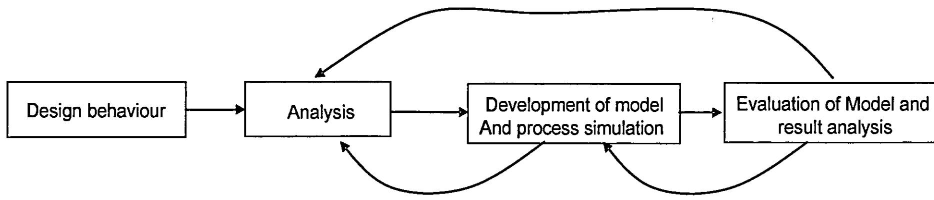


Figure 3.2 Iterative steps of this research

3.3 Validation framework

The importance of evaluation of design research is well recognised in the scientific community. Two important dimensions in evaluation are verification and validation. Verification deals with the truth or accuracy and the predictive power of theories, methods, and models, whereas validity deals with their relevance and meaningfulness (Pedersen et al. 2000). These are similar to the definitions that are presented in Table 2.1.

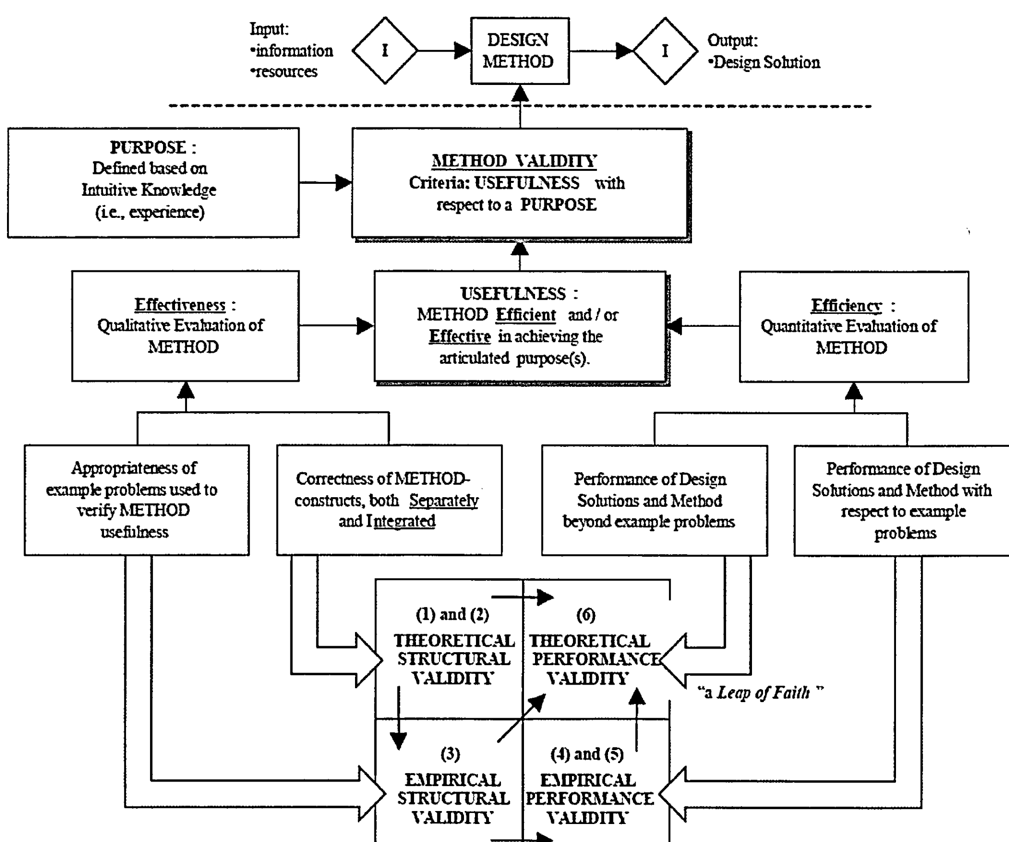


Figure 3.3 The validation square (Pedersen et al. 2000)

The following section describes a framework that has been followed in validating this research work. This validation framework is introduced by Pedersen et al. (2000).

Although this framework focused on the validation of "design methods" particularly, it is a valuable process for design research validation (Pedersen et al. 2000). They present that research validation is a process of building confidence in its usefulness with respect to a purpose. They associates usefulness of a design method with whether the method provides design solutions 'correctly' (effectiveness), and whether it provides 'correct' design solutions (efficiency). 'Correct' in this context, are design solutions with acceptable operational performance, which are designed and realised with less cost and/or in less time. Their validation process, supported by the validation square framework, aims at evaluating the effectiveness and efficiency of the design research, based on qualitative and quantitative measures.

Figure 3.3 shows the "Validation Square" (the square at the bottom), which consists of four validation squares, is a combination of theoretical and empirical aspects with structural and performance validations of the research results.

In the assessment of *theoretical structural validity* two issues are considered: (1) accepting the construct's validity, and (2) accepting method consistency. In order to build confidence in the validity of the initial construct constituting the method, they suggest using the literature. Well established and accepted references can build the confidence on initial construct and acceptance regarding the validity of an individual construct. The consistency of the design method, that is, how the constructs are put together, can be achieved using flowchart representations focusing on information flow. Therefore, it will be possible to demonstrate that for each step there are: (a) adequate inputs available, (b) that the anticipated output from the step (construct) is likely to occur, and (c) anticipated output is an adequate input to next step (construct).

The *empirical structural validity* - deals with the structural soundness for some particular instances, is concerned with (3) building confidence in the appropriateness of the example problems chosen for verifying the method performance. These can be addressed in three steps, by documenting, first, that the example problems are similar to the problems for which the method constructs are generally accepted, second, that the example problems represent the actual problem for which the method is intended and finally that the data associated with the example problems can support a conclusion.

In the assessment of *empirical performance validity* the issues considered are (4) accepting the outcome of the method is useful with respect to the initial purpose for some chosen example problem and (5) accepting that the achieved usefulness is linked to applying the design method. The usefulness of the method can be demonstrated through an example problem, and outcome of the method can be evaluated by linking metrics for usefulness to measure the degree to which an articulated purpose has been achieved. The purposes of a design method vary in an industrial perspective and a scientific perspective. From industrial perspective, the purpose might be linked to reducing the cost and time and improving quality, whereas, the purpose can be producing more scientific knowledge, from academic perspective. It also needs to be demonstrated that these achievements are linked to the application of the design method. The can be achieved by comparing the solutions with and without the method, allowing a qualitative evaluation.

The *theoretical performance validity* concerns (6) accepting the usefulness of the design method beyond the example problems. Confidence in generality can be achieved using induction rule that consist of the five steps from (1) to (5). Generalisation is attained through several case studies which assesses the theoretical proposition.

In Section 9.2, the validation of this study is presented.

3.4 Summary

This research is initially aimed at better understanding the role of testing in the product development process and how testing occurs in an industrial context. Through analysing the current state of testing and test planning in industry, this research also aimed to reveal issues and scope for improvement. The outcome of this research contributes to academic knowledge about testing in the product development process. Further, the research proposes methods to improve test planning in industry.

Chapter 4 Testing in an industrial practice

The main purpose of the cases study was to provide an understanding to the company's testing processes, but it was also carried out to identify the key issues that the company is facing. This study gives a good insight into current testing practice in engineering in a leading UK based company; which has not previously been comprehensively examined in academic research. A simple comparison study was also essential to determine if the issues highlighted in the company were existed in other companies.

This chapter first describes the case study company and its current testing practice. Next, the findings from this study are illustrated. The comparison with another company is presented in the following section.

4.1 The company, its product and product development process

The case study company is one of the leading suppliers of diesel and gas engines. The company was established eighty years ago and since then has produced over 20 million engines. It is a UK based company, over the years owned by many companies but currently owned by a leading USA company. The company offers a wide range of diesel and gas engines and power packages from 8.2 kW to 1886kW and has the capacity to produce up to 800,000 units per year. Around 30% of the sales volume is sold to the owning company in USA. The largest portion of sales is into the European market and the sales in the Asian market are rapidly increasing. Regulations and requirements in these markets vary widely. In these markets products are sold direct to original equipment manufacturers (OEMs) and distributors. Global product support is provided by 3,500 distribution, parts and service centres situated around the world.

4.1.1 Product

The company produces a wide range of engines. There are product families with different power ranges to meet the requirements from different markets. Products also vary in families depending on the number of cylinders, aspiration and control mechanism. Figure 4.1 shows the series of engines in the company's product range.

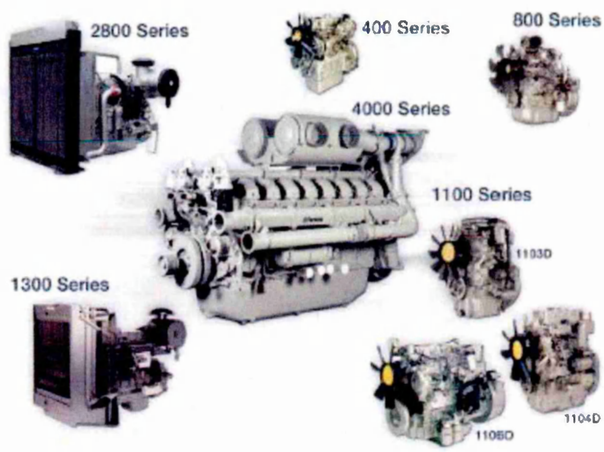


Figure 4.1 Company's product range (taken from company's presentation slide)

These engines are used in many applications such as agriculture, construction, material handling, marine, general industrial and electric power. Figure 4.2 shows a pie chart indicating the percentage of products sold for different application in 2006.

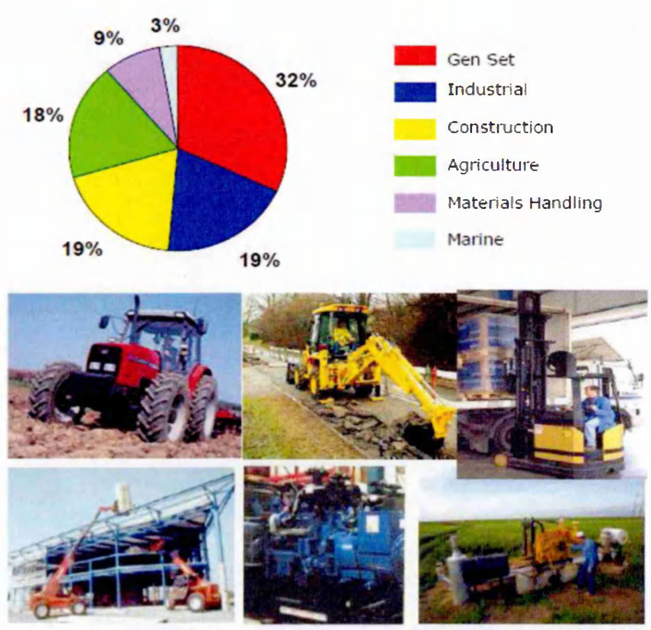


Figure 4.2 Company's product application, pie chart shows the percentage of product sold in different applications in 2006 (taken from company's presentation slides)

These engines are also used in a wide range of environments. Environmental conditions like extreme cold or hot weather, or physical machine working such as a desert, where sand can enter the machines causing problems. Adapting engines to fit different application environments is a standard activity for the company.

4.1.1.1 Design challenges

A key challenge for the company is to comply with increasingly tough environmental legislation. This has led to huge technological changes accompanied by a significant decrease in product development time. Various aspects of engine performance are being regulated; the most important area is that of exhaust emissions, particularly, exhaust gas emissions compared to emission standard Tier 3/Stage IIIA levels. Tier 3 is an EPA (the US Environmental Protection Agency) terminology and Stage IIIA is the corresponding EU terminology. Tier 4 Interim/Stage IIIB standards which have been in effect from 2011 require a 90% reduction in PM (particulate matter) and a 50% decrease in NOx (Oxide of Nitrogen) emissions in the U.S. and Europe. Tier 4 Final/Stage IV standards, which become effective in 2014, specify NOx reductions by an additional 80%, taking PM and NOx emissions to near-zero levels. Additional information about these regulations can be found in (EPA's Office 2013). These requirements for PM and NOx emissions can vary significantly for different parts of the world.

4.1.1.2 Product development time

With the introduction of new emission standards, the product lifecycle time for development and manufacturing of engines has been reduced dramatically. Because, of the new legislation, current product becomes illegal in many markets and new products must be launched into the market and older ones withdrawn. Jarratt (2004) compares the length of time in production for four recent 4-cylinder engines of approximately 1 litre per cylinder capacity (Figure 4.3) The 4.236 engine was manufactured for approximately 30 years, which is five times as long as the 1104 engine.

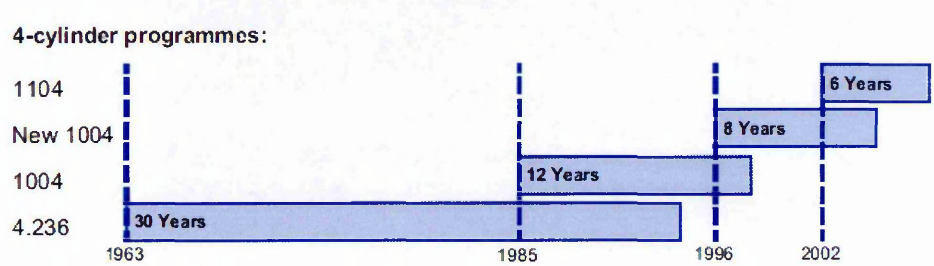


Figure 4.3 Decrease in engine life cycle time (Jarratt 2004)

With the increasing pressure to meet the emission legislation, the priorities of product development have shifted significantly. Previously, engineers would focus on improving the performance of an engine, such as power vs speed or durability. Now, engines must comply with the emission legislation to be qualified to sell, although, factors like fuel economy, reliability and product support are also vital for survival in the market.

Legislation is a major influence on technology development. As Engineer 1 commented "*legislation is driving the technology*". New technologies, such as precise fuel injection techniques, or after treatment equipment, are integrated with the engine system. This increases the complexity of the product. The number of electronic components has been increased for control and measurement with recent engines having an Engine Control Module (ECM). These extra components and subsystems have increased the complexity of the product as well as amplifying the effort, time and cost required to build, test and manufacture the product. As a result the cost of projects is growing in each generation of engine build and manufacture.

4.1.2 Product innovation

Although, engineers did not mention product innovation specifically, the analysis of the interviews identifies that product innovation in the company happens in the two dimensions of "product capability" and "technological capability"(Veryzer 1998). Product capability refers to the benefits that are perceived and realised by customers or users. Product capability is realised by the customers/users through the functional capability of the product. The technological capability refers to the way that a functional capability of the product is achieved. It is often behind the scene and its full extent is not realised by the customer. Products may be perceived as being essentially the same, as previously existing products, even though they utilise new technology. The diesel engine is increasingly coupled with after-treatment devices, such as, particulate filters and NOx catalysts, which might have little impact on consumers, in terms of product benefits or use, like, fuel consumption and durability, which have not improved.

The engineers in the interviews describe their products as incrementally improved products. They refer to the fact that the "functional capability" of the diesel engine doesn't improve dramatically but that frequent incremental changes in the technology improves performance or lowers cost of ownership.

The company's innovation is measured in terms of the "degree of changes" which happens between products. This is similar to the definition of Wheelwright (2010) who suggests that the degree of change, inclusive of the degree of manufacturing process change, is the most useful way to classify development projects.

When mentioning the natures of testing for different types of product innovation, Engineer 1 mentioned,

> "incremental design changes can be managed with standard testing or validated simulations whereas step changes (radical product innovation) may require a new test plan".

For example, if the company is designing a cylinder block which is a scaled version of a previous product, critically stressed areas would be already known. Thus it might be possible to assess the risk accurately through simulation, whereas a new cylinder block will be physically tested.

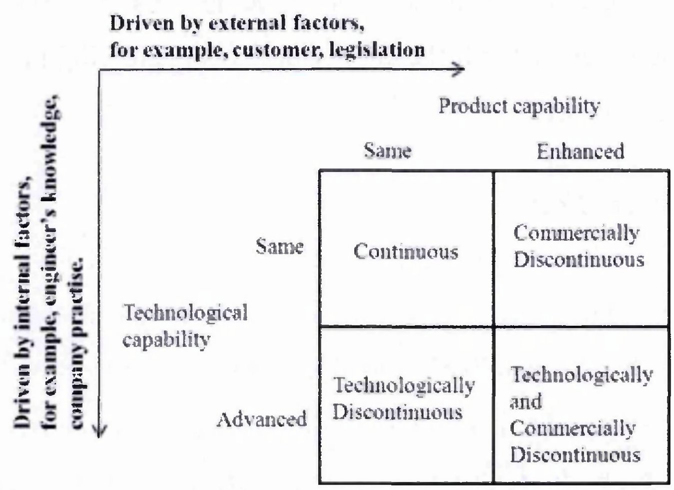


Figure 4.4 Different factors that are affecting the "product capability" and "technological capability" dimensions (overlaid on diagram from Veryzer (1998))

The dimensions "product capability" and "technological capability" are driven by different factors in the company. These drivers are shown in Figure 4.4, overlaid on the diagram of dimensions of product and technology capabilities referred to in Figure 2.2 of Chapter 2, derived from (Veryzer 1998). Customers mostly drive improvements or innovations in product capability dimensions. For example, the company continuously improves its engine to achieve desired requirements on power, emissions composition, and fuel consumption derived from customers' wants. On the other hand, the technological capability is the way the product's functions are performed, and is mostly driven by the company engineers' technological understanding and expertise being brought to bear as particular technologies are assessed for inclusion.

4.1.3 Product development process (PDP) and testing

A company's new product introduction (NPI) process is a systematic planning and control process, which includes all activities to develop and introduce new products to the market. A company will use their generic product development process minor modifications, which depend on the particular type of product.

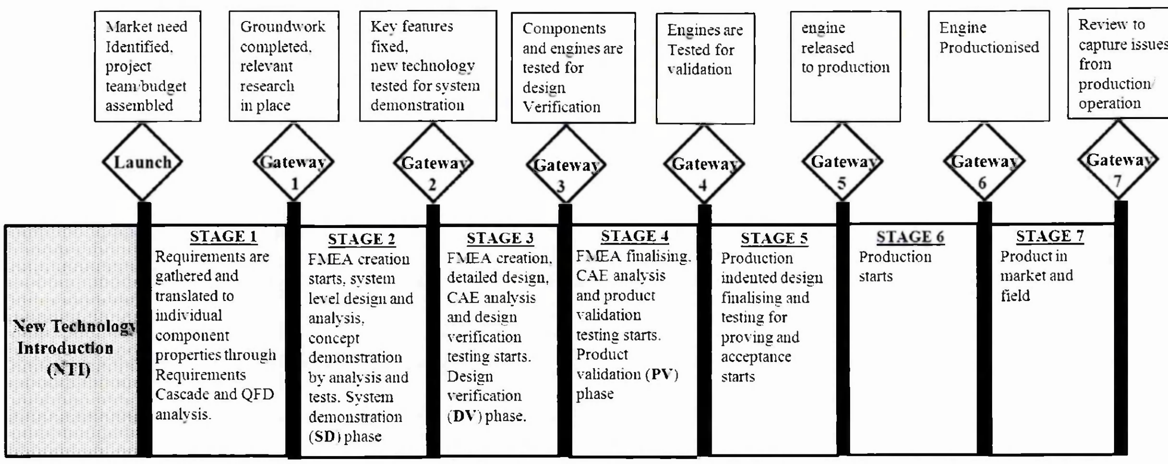


Figure 4.5 An outline of the company's Stage-gate NPI process

The case study company has a structured gateway process for New Product Introduction (NPI) that has seven stages (see Figure 4.5). Each stage leads to a formal gate review, starting from "Launch" to "Gateway 7(GW7)". Based on prescribed criteria, a product must pass through gate review before the product development project proceeds to the next stage. The case study company also has a New Technology Induction (NTI) process which occurs in a different department. This NTI takes place as a general research and development exercise, before the NPI process starts.

Most of the testing occurs between stages 2 to stage 4. This means that testing happens from Gateway 1 (GW1) to Gateway 4 (GW4), and often until Gateway 5 (GW5). Development testing starts between GW1 to GW2 when the technology has been identified and continues till GW4, after which the engine is released to production. This thesis research focuses on these three main phases of the PD process, namely stages 2, 3 and 4. Testing in the company falls into three stages and serves a different purpose in each stage:

(i) Concept/System Demonstration (SD) in stage 2

(ii) Design Verification (DV) in stage 3

(iii) Product Validation (PV) in stage 4

In each stage, Performance and Emission (P&E) testing starts first and then mechanical durability and reliability are tested. Each of these phases is now described in turn.

Concept/system demonstration (SD) testing is primarily to demonstrate 'performance capability'. It shows that the technology can deliver the required performance. Alternative concepts are analysed and evaluated. A combination of old and new parts are built into an engine called a MULE. This MULE engine is tested to verify the performance of new parts. As new parts arrive the old parts are replaced and testing continues. The product specifications evolve as more design decisions are taken during this phase. It is assumed that by Gateway2 (GW2), the concept will be selected, the component will be specified and the whole engine will be built with at-least some production parts, will be ready to be tested for Design Verification (DV).

Design verification (DV) is primarily to develop optimal performance and validate hardware at the optimised performance. The aim is to ensure that design outputs meet the given requirements under different use conditions. At this stage, testing focuses on the verification of a chosen design, through detailed analysis and testing of stress, strength, heat transfer and thermodynamics etc. This stage validates the hardware prior to commitment to expensive production tooling.

Product validation (PV) checks the effect of production variability on performance and any remaining hardware variation. This phase performs hardware testing which is limited to late design changes and emissions conformance testing. In this phase, detailed testing for reliability and durability occurs and the intended product is validated. The mandatory tests required for compliance usually occur during PV phases.

In Figure 4.6, a flow diagram of testing and related activities is presented. It is derived from information obtained in the company interviews. Any component level testing happens individually when engineers want to investigate a concern; otherwise component level testing occurs at the component's suppliers. Engine level testing refers to standalone engines on a test bed. Machine level testing refers to the case when engines are mounted in machine/vehicles for testing under expected use conditions. Figure 4.6 illustrates how engine level and machine level testing are mainly conducted in parallel in the three stages of concept/system demonstration (SD), design verification (DV) and product validation (PV).

Design, CAE, Procurement and testing activities undergo at least three iterations from Stage 2 to Stage 4, as shown in Figure 4.6. At each stage, the engine level testing contains a large number of tests. Some tests are grouped and some are individually conducted. The company's product development process is heavily involved in testing activities from the start of the project.

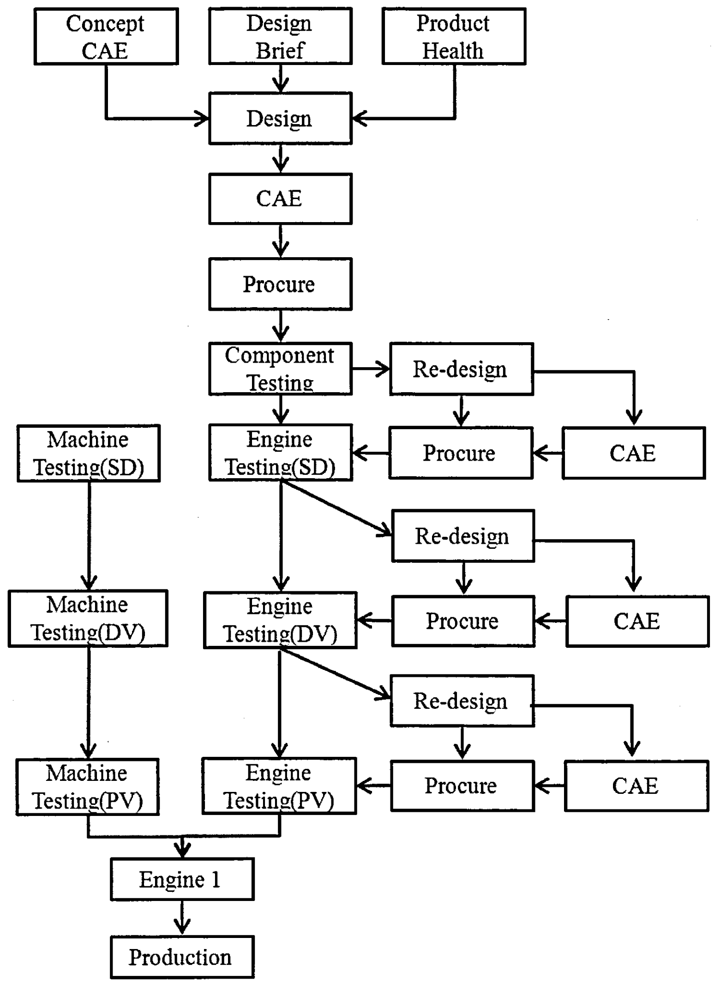


Figure 4.6 Flow diagram of testing and related activities

4.2 Testing planning

In this section, the company's process of testing planning is presented, with associated tools that are used in this process described in the next section, with an analysis of why and how these tools are used during the process of testing planning.

The basic structure of testing activities is planned before GW2. A flow diagram of the early validation and testing plan is shown in Figure 4.7.

Initially, potential failure modes of design are analysed by assessing all the possible ways a design can fail to meet its requirements. All components, subsystems and the engine are assessed individually to identify potential failure and risks of these failures. Health monitoring data from the previous engine and characteristics of product and critical customer requirements are used as input to focus on identifying potential failures. Potential failure modes are identified so that design actions can be taken to prevent or minimise the effect of these potential failures.

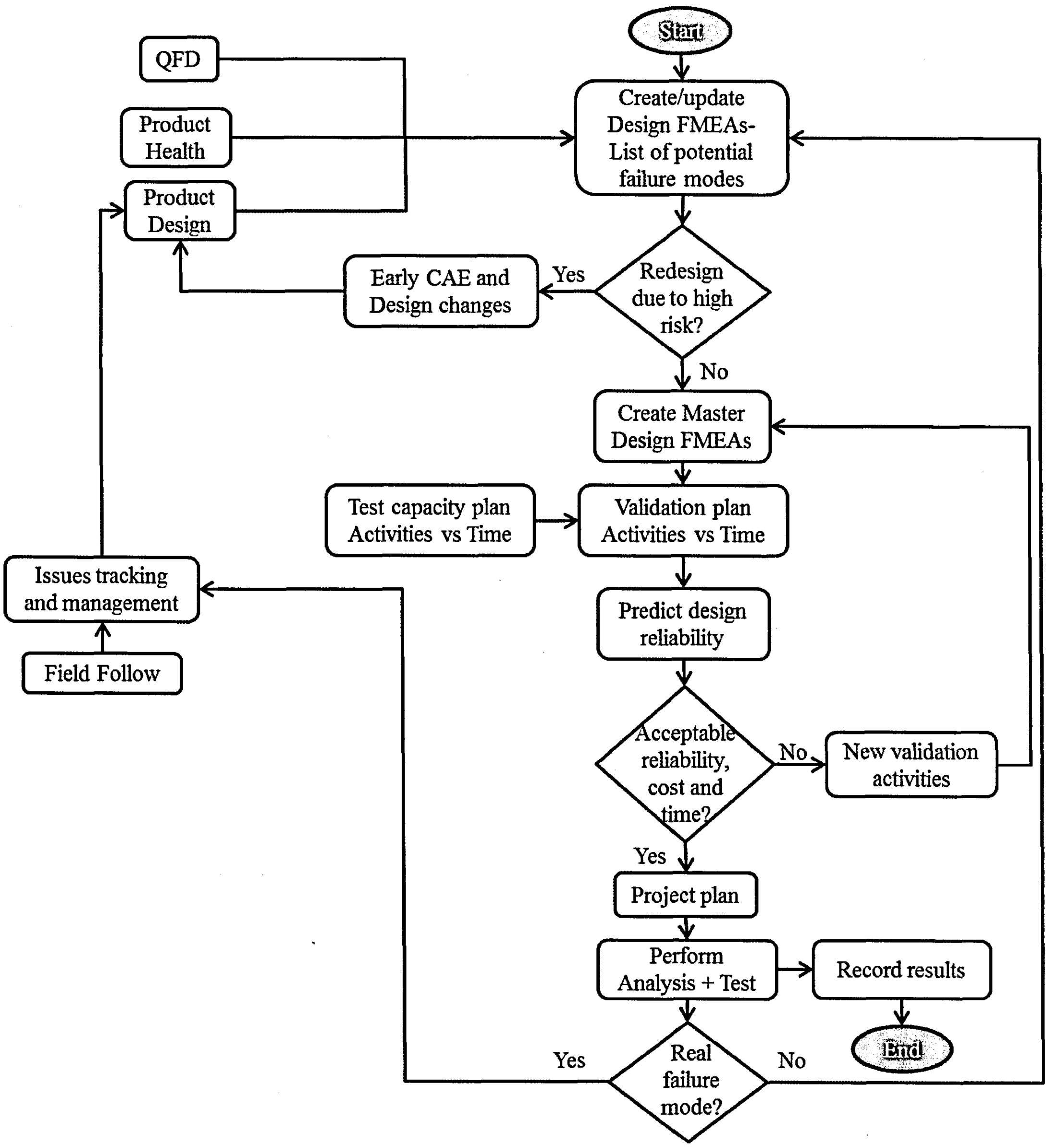


Figure 4.7 A schematic view of the case study company's iterative process of testing and validation planning.

As the diesel engine is a mature product and design changes happen incrementally in the company, engineers start with an existing analysis of the previous generation of products. For a new product introduction (NPI) programme, product objectives are checked against a current product issues (CPI) database. The CPI database provides information about failure modes and effects of current products, which will need special attention for next generation of product. This process is carried out by key team members who are the technical specialists, component owners, design owners and the verification and validation managers.

This process starts between GW1 and GW2, i.e. before the concept demonstration phase. Initially, the design alternatives are included in the analysis, because selection of a design is made based on the risk with that particular design and the associated time and cost of its validation program. All design options are considered during the early stages of the product development planning. These help to analyse the trade-offs that can be made across different design options. If this analysis identifies high risks in design decisions, further CAE analysis and design changes are planned to be performed until the risk is reduced to an acceptable level to proceed with the project. When the overall risks are assessed, design verification and validation actions are decided and planned to mitigate risk. Verification and validation activities can range from design changes, CAE analysis to testing.

A validation plan needs to be adjusted to fit the planned use of the testing facility. As the test facility is shared by different projects and programmes that run in parallel. Different projects request to book the testing resources- facilities and staff. Therefore, a validation plan can be significantly affected by the existing availability and schedules of the testing resources as well as the priorities assigned to projects and customers. When time is an issue, urgent programs can get the highest priority.

Decisions are then taken around how much the risk can be reduced by a testing and validation programme and how long and how much it is going to cost. A validation plan must not only reduce the technical risks or uncertainties, but also needs to be within an affordable time and cost to the project. Different or alternative validation activities are requested when the validation plan costs beyond the budget. Finally, the validation plans together with the test facility plan are placed in overall project plan.

At GW3, engineers aim to finalise the tests and planning for product validation. However, the development of a validation plan as shown in Figure 4.7 is a dynamic process, continuing until GW4. Every time a validation activity takes place, risks are assumed to be reduced. A key evaluation measure for a testing plan is the expected risk reduction that it will effect.

There are three main decision points in this process:

(1) risks are assessed initially to decide if an engine programme is viable to proceed,

(2) a validation programme is checked for feasibility in terms of the cost and time it requires and reliability it can achieve, and

(3) a real failure is recognised through a validation activity, especially during a test.

A failure in a test, especially towards the later stages of the programme, can have a huge impact. It can destabilise product development requiring design changes and subsequent re-planning of the validation activities.

4.3 Tools for testing planning

There are mainly three tools: Quality Functional Deployment (QFD), Failure Modes and Effects Analysis (FMEA) and Design Verification Plan and report (DVP&R) that were mentioned by the engineers during the interviews. QFDs are used to translate the customer requirements to products technical requirements, FMEAs are used for risk assessments of these technical requirements and DVP&R for validation planning. Figure 4.8 shows a schematic of data flow between these tools. Ideally QFDs drive the technical requirements of FMEAs and FMEAs drive the testing activities, which are planned in DVP&R. How these tools are used in the company is described below.

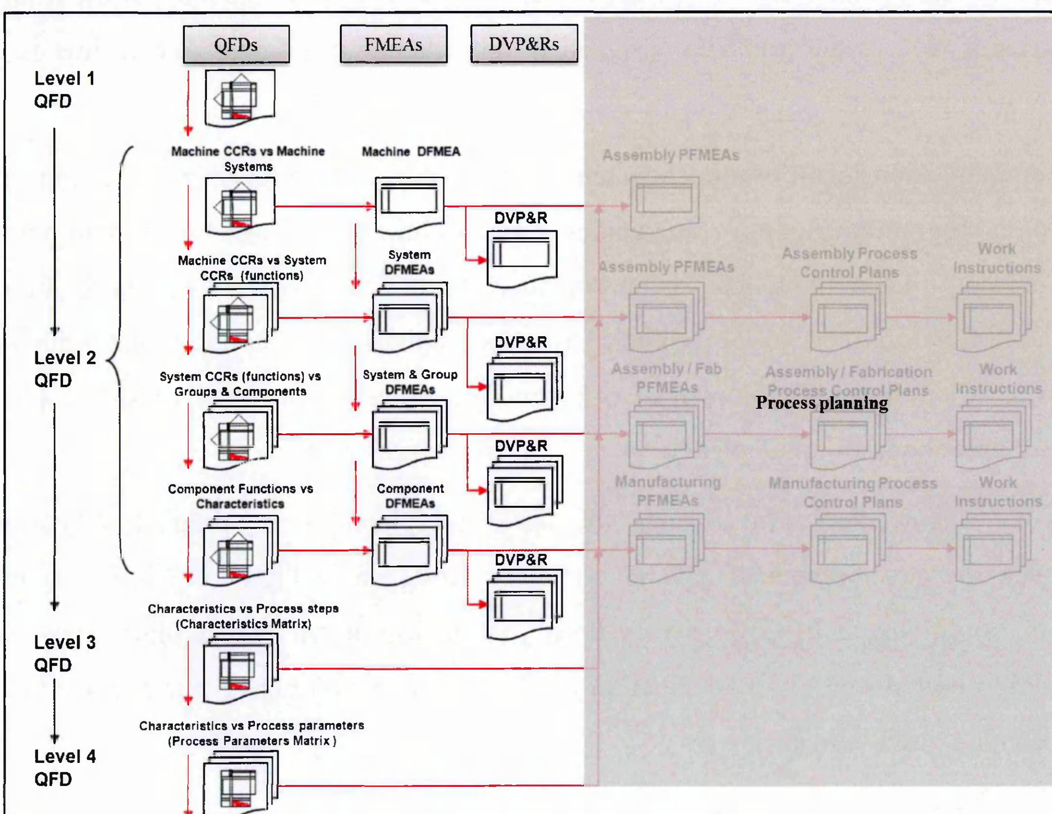


Figure 4.8 Data flow from QFD to FMEA to DVP&R in the case study company

The FMEA tool has been introduced into the company by the parent company. The company follows the ISO FMEA format and best practice guide to build their FMEAs. As this tool has been used for more than twenty years, it has a mature level of application.

Engineer 3 commented that

"FMEA is the driving tool that is directly used for test planning"

On the other hand, the company has used the QFD tool for the last seven years, and is still exploring its capability. They realise that while there remain limitations in the effectiveness of their application of QFD, FMEA is the main driver for the testing planning. The following subsections describe the ways that these three methods are used in the case study company are now described in the following subsections.

4.3.1 QFD

QFD utilises the house of quality (HOQ), which is a matrix (as shown in Figure 4.9) providing a conceptual map for understanding Critical Customer Requirements (CCRs) and establishing priorities of Technical Objectives (TOs) to satisfy them. An HOQ typically contains information on CCRs, including the customer importance of the individual CCRs. There are also TOs for satisfying the CCRs, which effectively describe the relationships between CCRs and TOs.

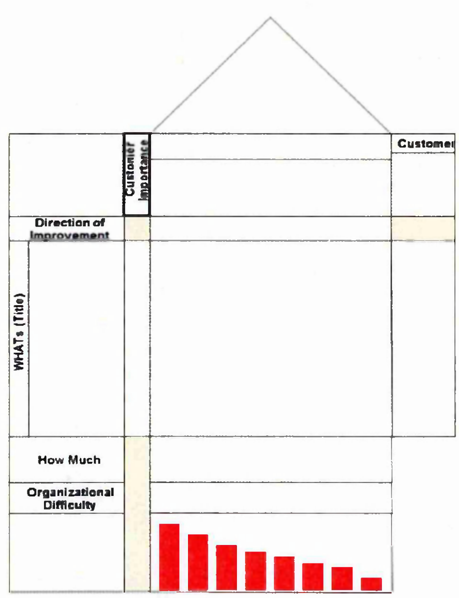


Figure 4.9 A template for QFD House of Quality

In the company, 'voice of the customer' (denoted VOC in Figure 4.10) is captured in many ways: directly, through discussions, interviews and workshops with customers, and indirectly through analysing customer specifications, warranty data, and field reports etc. and through dealers and distributor channels. As one engine may be used in different machines, the engine maker may not have a direct connection with the end consumer, understanding the needs of consumer if often challenging. Also, since the company's NPI programmes run for 4 to 5 years, many of the current customer requirements are actually the predictions on what customer's might want in future.

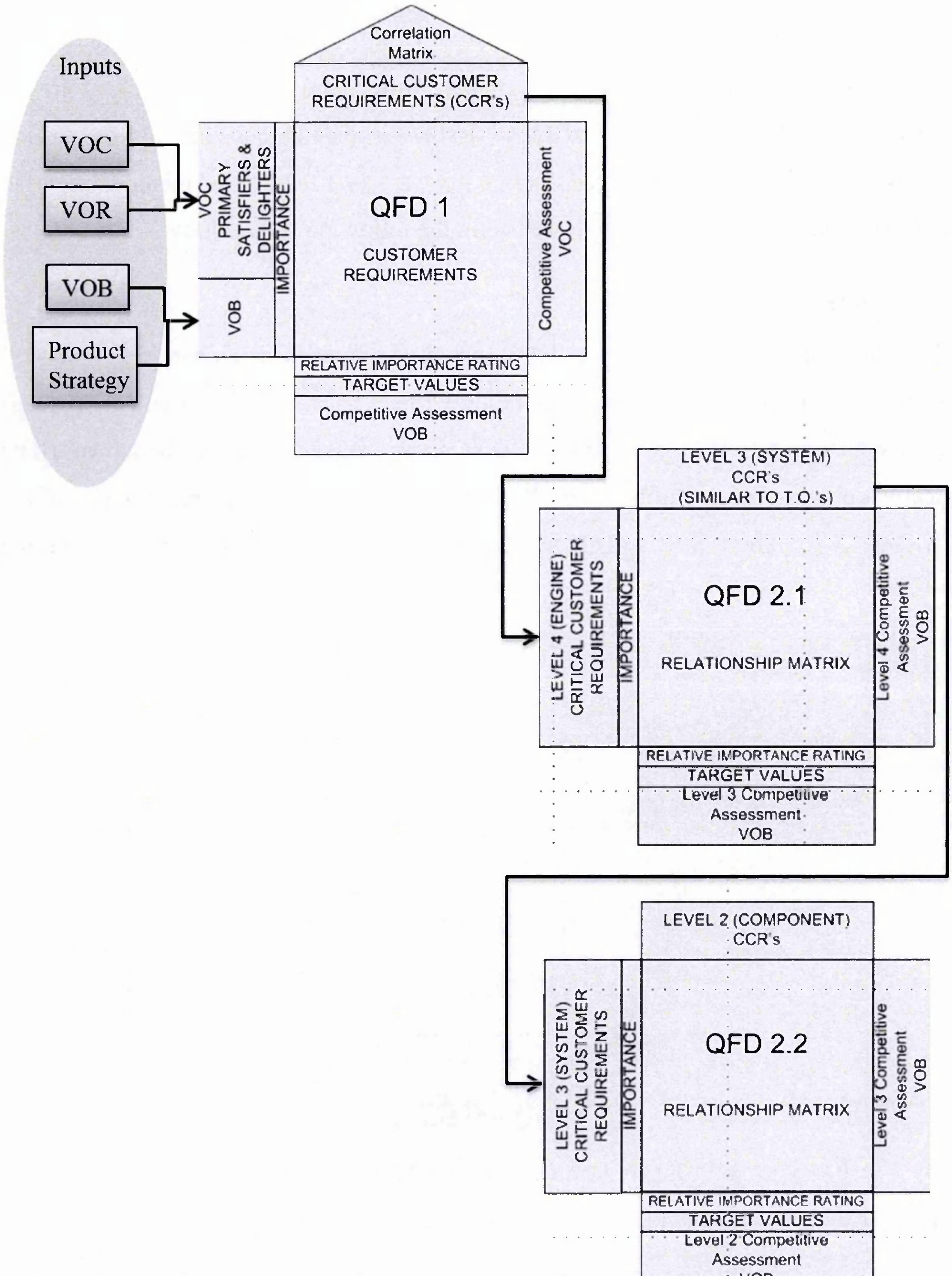


Figure 4.10 Two phases of QFDs and their flows

The needs of customers are diverse for this company in terms of product characteristics and conditions of use. These diverse voices are analysed to understand the important areas where negotiation is required on the conflicting customer requirements. In this way a balanced range of products targeted to specific customer groups are developed. Once critical customer requirements are identified, these are summarised in a "house of quality"(HoQ). 'Voice of regulations' (denoted as VOR in Figure 4.10), such as, emission control, are also considered as CCRs and listed in the HoQ. These matrices are used to translate higher level "what's" or needs into lower level "how's"- Technical Objectives (TOs) to satisfy these needs.

As mentioned in Section 2.5, a conventional QFD process has four phases: product planning, product design, process planning and process control planning. This case study company mainly uses the first two phases of product planning and product design. One or more matrices are prepared during each phase. These QFDs are represented in Figure 4.10. Product planning QFD, sometimes termed QFD1, is used to translate the overall customer requirements to engine level CCRs. Product design QFDs have two levels: 1) QFD 2.1 translates engine level CCRs to its sub-system level CCRs or Technical Objectives (TOs) and 2) QFD 2.2 translates sub-system level CCRs to component level CCRs or Technical Objectives (TOs).

The QFD process starts between Launch and Gateway1. Customer requirements can change through better understanding of the customer's needs or customers directly specify something different. Therefore the QFD process evolves with these changes. The company tends to be firm on these requirements after Gateway 3 where the products capability are realised and production ready designs are prepared.

4.3.2 FMEA

There are two types of FMEA analysis that are performed in the case study company: Design FMEAs (DFMEA) and Process FMEAs (PFMEA). DFMEAs are used to assess design risks, and risks in the processes of production and supply are assessed through PFMEAs. In this thesis, only DFMEAs are considered because the research focuses on product testing.

The creation of DFMEAs is a top-down approach (Figure 4.11). An Engine level DFMEA is built first, and then the sub-system level and component level are created successively. The engine level FMEA considers engine's overall functional and technical requirements. These requirements are gradually broken down into lower levels, i.e. subsystem (such as the cylinder head assembly), and component (such as piston) level requirements. These are assessed through lower-level FMEAs. For an engine programme, there can be more than fifty FMEAs.

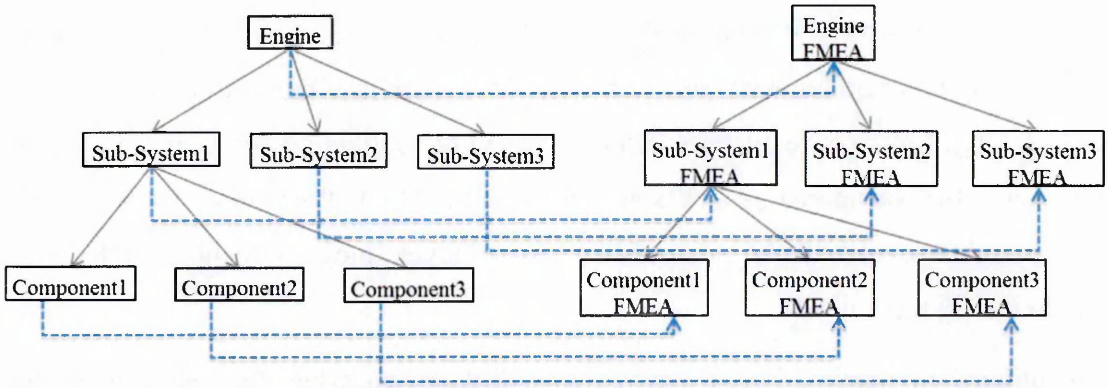


Figure 4.11 Engine decomposition (at left) and its links with FMEAs (at right)

DFMEA analysis includes the following steps: identifying possible failure modes (what could go wrong?), failure causes (why would the failure happen?) and analysis of failure effects (what would be the consequences of each failure?).

4.3.2.1 The formation of a DFMEA

For a new product introduction (NPI), the output from the QFD process, namely the customer requirements, product objectives, technical objectives, and conceptual design are all considered in assessing the product's feasibility. The formation of the FMEAs start after the feasibility of the concepts are analysed and chosen.

Initially the process of building an FMEA starts with a cascade from the engine level, combined with the definition of what the engine should be able to do to meet the design objectives. The engine FMEA includes the failure modes (the ways that an engine might fail to achieve requirements), corresponding failure causes and failure effects. From there, system level FMEAs breaks down the engine into core individual systems, for example, combustion system, after-treatment system, etc. These FMEAs consider what a subsystem needs to achieve to deliver the overall engine performance.

Component level FMEAs consider what each component needs to achieve a system level performance. The component level FMEAs tend to analyse individual ways a component can fail. Many of the detailed failure mode analyses occur at component level FMEAs. The system levels FMEAs have more emphasis on failure to meet performance criteria or failing to meet durability criteria. Some important characteristics of the product like reliability and robustness are also included in FMEAs as functional requirements. For example when the company is designing the products to last for X thousands hours specified under a profile of use conditions, they build this into a functional requirement. The company translates non-functional requirements into functional requirements and puts these requirements into FMEAs.

FMEA creation starts with a list of functional requirements (FR) of the product under analysis. For a clear and focused FR description, it is considered essential that this is driven by the engine FMEAs. In that way, product objectives can be built into the engine FMEAs, which are cascaded into the sub-system FMEAs, and then into component FMEAs. Therefore the FR can be realised and quality of the FR can be improved. Also, well-defined product objectives and technical objectives help to build up clearer functional requirements, as Engineer 3 mentioned:

> "more experienced people can put the functional requirements and failure modes more clearly than the people with less experience. Less experience people tend to describe the symptoms rather than the functional requirements".

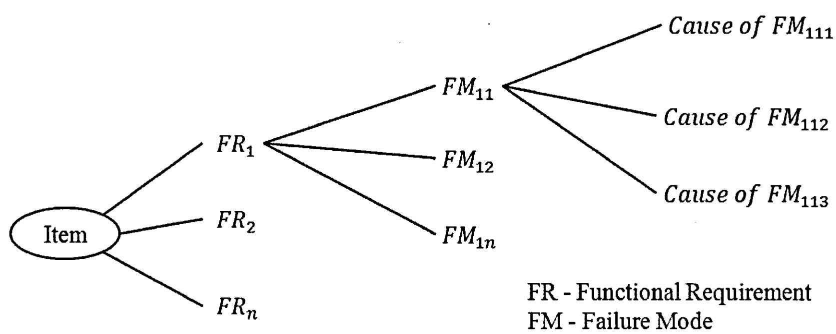


Figure 4.12 A functional requirement can have several failure modes and there can be several causes for a failure mode

Each of the FRs can fail in several ways; therefore can have multiple failure modes. Also there can be several causes of each failure mode. So, all the possible causes of failure are listed against each of the failure mode (as shown Figure 4.12). Each of these causes is then scored against the probability of occurrence of that failure mode. Occurrence rate is established on the data from failure logs, warranty data, maintenance logs, service data etc. The company uses an Excel spread-sheet to maintain the FMEA analysis. A snapshot of an example FMEA is shown in Figure 4.13.

At a component level, the effect of failure is analysed in terms of the effect of failing to achieve a functional requirement of that component; these are the primary effect of failure but also secondary effect on upper levels, are considered. A potential effect of failure is a failure at the next higher level design, perhaps system or engine levels. Effects analyses of failure modes are usually bottom-up approach. An example shown in Figure 4.14, considers hardware through an FMEA approach. However, the effects analyses could cover hardware, functions, interfaces or a combination of all these items.

Item	Functional Requirements	Failure Mode	Cause of Failure	O	Primary Effect of Failure	S	Secondary Effect	S
Dipstick (blade or spring) & tube assembly	Measure sump oil level	Does not measure sump oil level	Marking unable to read	3	Incorrect oil levels	6	Engine failure	8
			Dipstick too tight in tube	5				
			Dipstick snags on tube	2				
			Unable to replace dipstick	5				

Figure 4.13 An example of a FMEA analysis on a spread-sheet of a dipstick (only a selection of columns in the spread-sheet is shown)

To identify the potential effects, the company will also review documents, for example historical data, warranty documents, field service data, customers complains etc. depending on the seriousness of the effect of a failure mode, company rates the severity. The primary or secondary effects together determine the severity score.

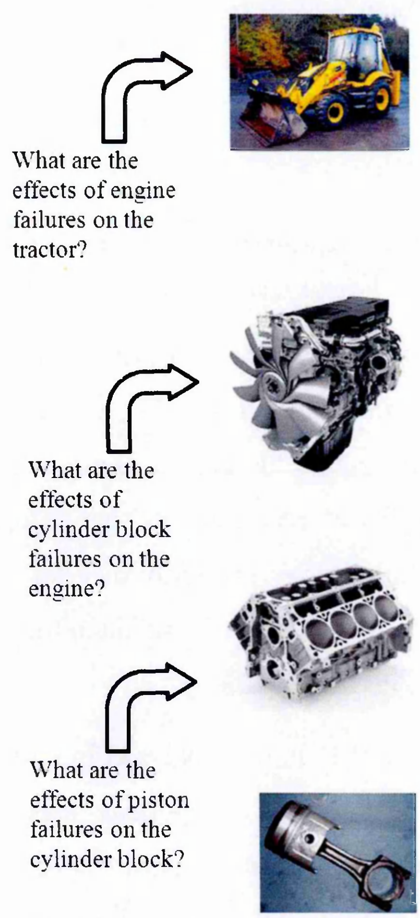


Figure 4.14 A bottom-up approach to analysing the effects of a failure of a component

The scoring of the FMEA for an existing component/product is an evolving process. Initially the score is an assumption rather than a real value based on current evidence. For a new part or product initially assumptions are made based on the previous knowledge.

For a completely new component or for new application these scores are set to relatively very high initially. The FMEA is not a tool with the capability of indicating the newness in a visual way. Just assigning a higher score on a new functional requirement doesn't differentiate it from the functional requirement that had the same high score because of seen failure. The company doesn't identify (initially) any qualitative difference between new functional requirement and a functional requirement with previous failure history. Both of the scenarios will get the same attention for risk minimisation.

All the failure modes are scored against the occurrence and severity. Risks are calculated. One way that this is done in the company is through a Risk Priority Number (RPN) technique for analysing the risk associated with a potential failure. However, this company uses a more comprehensive method. Figure 4.15 shows a metric which uses severity and occurrence as measure of risk. It maps out particular cases where risk mitigation and where the risks are too great to be mitigated. Risk mitigation activities are allocated to reduce risk below threshold levels. The scores are revised at each step of the programme when a realistic insight takes place through design analysis or test.

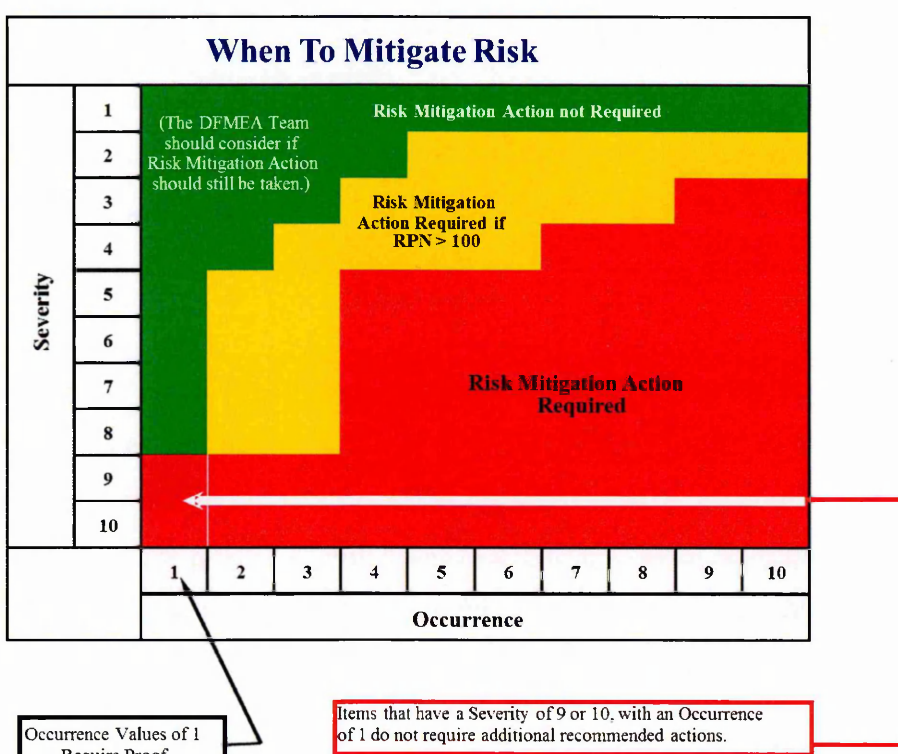


Figure 4.15 Types of risk mitigation measured along the axes of Severity and Occurrence

Design and CAE activities and associated resources are assigned to each of the failure modes to reduce the risks to acceptable levels. Then validation activities and resources are required to prove that the risks have been mitigated. Some tests are performed to justify whether an estimated risk is a true risk and whether the risk has been over or under estimated.

After building the FMEA, the development team start planning actions to mitigate the risk. The recommended course of action adopted in the company is to design the critical failure areas first. Critical areas are colour-coded. Monthly reviews determine which FMEA contains high risk items that have not had actions taken on them. This reminds engineers to request simulation as well as developing and analysing the current design.

The cost of achieving a requirement is not considered in an FMEA. The cost is kept at the technical level and dealt through the NPI process. It is the Process FMEA rather than the Design FMEA that is more concerned with time and cost. However, cost and time do need to be linked up with the design. For example, whether to choose an expensive material or component for testing would not be a part of the Design FMEA, but would be part of the discussion around the validation plan where the company looks for cheaper alternatives which can be validated.

FMEA ranks potential failure modes based on three criteria: frequency of occurrence, severity of failure, and difficulty of detection. Each failure mode is assigned a score for each of these three criteria on an ordinal scale, where a larger score indicates a less desirable circumstance. Criticality matrices (Figure 4.15) highlight the potential risks associated with the failure modes of each functional requirement. The FRs with scores (occurrence) above 3 trigger actions through additional design or simulation or by testing.

New functional requirements will tend to score relatively highly as newness is associated with perception of high risk. The FRs for which there is failure evidence might be equally severe as the new FR but perceived as lower risk. The performance against new FRs can still be improved through design and simulation whereas the company have tried to design out failure against the previous FR but have not succeeded. Different actions are required for these two types of FR.

4.3.2.2 From FMEA to validation plan

At the beginning of a project, the company sets acceptable risks related to the technology available at the time for the project. Through the development process, engineers aim to bring down any risk below a company set threshold. FMEA tools help to highlight the critical areas based on initial risk analysis. Engineers assign design changes, CAE analysis and tests to mitigate the risks. These actions need to be proved to minimize risk and are used to formulate the validation plan. A validation plan contains not only validation activities and tests, but also the upstream CAE analysis and verification activities.

Critical areas highlighted in FMEA are dealt with further actions of CAE analysis and design changes. For an NPI programme, risks of a failure mode may not be totally understood, so some baseline testing is performed to investigate the true mode of failure or to rank the failure mode correctly. Some of the failure modes will be investigated through testing. After taking any planned actions, these risks are assumed to be minimised and the design will be subject to formal verification and validation. Figure 4.16 shows the typical steps of this FMEA-based design and testing process.

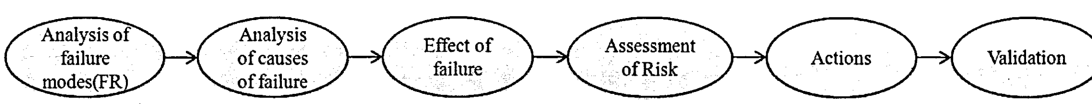


Figure 4.16 Steps of FMEA analysis, design actions and validation

FMEA owners make constant reference to system boundary diagrams and technical objectives. They are concerned about the interaction between systems and in finding the potential failure modes that might not appear at the top of a critical list. Two levels of FMEA outputs emerge. The first highlights critical component and system issues while the second uncovers interactions and hidden critical issues.

Particular attention is required when system level FMEAs are formulated, because this is when the interactions between system and component level FMEAs emerge. If the system levels FMEAs are isolated, required design changes can get overlooked. Therefore system level FMEAs are cascaded to limit these issues. However, this cascading also requires spreading the ownership of the component/systems FMEAs and that can cause a problem. To counter this, the company have introduced new teams for CAE system, performance software and cascade and mechanical durability and reliability, who are responsible for maintaining these interactions of the FMEAs over and above the individual FMEA groups.

Usually component owners, design engineers and CAE engineers can request for tests if they feel a particular area needs further understanding through testing. Therefore all the requested tests are assigned with particular failure modes. So the FMEA based tool is capable of producing a list of failure modes that are linked with requested tests. This list tells the validation team what failure modes they are actually going to be signing off and indicates the validity of the test against associated failure modes.

The verification and validation managers are responsible for accepting and rejecting a test. Sometimes they group tests together while in other cases they disaggregate the tests. The validation managers decide what do they need to do with the engine and design the tests to prove those failure modes. For example, if a failure mode relates to wear issues, they will need to plan to measure before and after the test. If it is a performance related failure mode, they might instrument the test to measure how the parameters are performing throughout the test. They also get an indicator of the risk levels of the failure modes, which helps them to justify the risk reduction level against the target. Once the risk mitigation activities are accepted by the validation team, they start to link up the failure modes with an approved test. On the other hand there can be several reasons for rejection of the risk mitigation. For example, the test might not be realistic to perform, or they have already done it, it may be too complex, or it may be expressed confusingly and badly worded.

The validation managers also plan an appropriate sequence of test work considering the stages of the Gateway (GW) process. To do so they put the dates against each of the activities. The main purpose for the validation team is to plan test activities, decide what test work needs doing, when to start and when to finish. The team also monitors in each stage of the programme how these risks are being reduced. If the risks are not reduced quickly enough then the sequence of specified test work is reconsidered to identify whether the appropriate tests are being carried out against the time and cost constraints of the programme.

There are two levels of planning. The first is planning the tests arising from the FMEAs. This is essentially a validation or verification plan. The second is project planning which centres around Gateway targets, costs and times. Both of these levels of planning finally contribute to the test plan. The first validation and verification plan needs to be implemented in the overall project plan. These two plans interact and require careful capacity management in terms of time, cost and availability of the test facilities.

4.3.3 DVP&R

DVP&R stands for Design Verification Plan and Report. The creation of the DVP&R is an integral part of the FMEA process. The FMEA process has four steps. In Step 1 the potential risks are defined. In Step 2 design actions proactively mitigate the highest risks.

In Step 3 the analysis, simulation and test actions are planned to demonstrate the real extent of the risks. In Step 4 the plan is implemented and learning is fed back into the FMEA. Each FMEA has its own DVP&R. Many of the actions listed in one FMEA will also be referenced in many other FMEAs.

DVP&R encompasses all aspects of design verification, from initial CAE analysis to hardware testing. This tool helps to analyse both the internal and a supplier's design verifications (DV) and product validation (PV) plans. The focus of DVP&R is the verification that the DV/PV process is being carried out effectively, thus ultimately ensuring product and process quality. A DVP&R lists each individual test and CAE analysis, when it was performed, the specifications, acceptance criteria, results and the assessment pass/fail (see Figure 4.17). Engineers in the Case Study company use the terms DVP&R and validation plan interchangeably. The 'validation plan' should in principle be a subset of the DVP&R. Further, a DVP&R created after Gateway 4 is essentially a validation plan.

Company name	Design Verification Plan and Report			
Prototype:	Prelaunch:	Production:	Key Contact/Phone:	Created:	Modified:
Control Plan Number:			Core Team:	Customer Engineering Approval/Date (if req'd):
Part number/Latest Change Level:			Supplier/Plant Approval/Date:	Customer Quality Approval/Date (if req'd):
Part Name/Description:			Date/Other Approval (if req'd):	Date/Other Approval (if req'd):
Supplier/Plant:	Supplier Code:			

Test Number	Test Name	Test Method	Acceptance Criteria	Test Location	Sample Size	Start	End	Assigned To	Report Number	Status	Start	End	Sample Size	Test Results	Completed By	Notes (Results)

Figure 4.17 A template for DVP&R used in the case study company

The Design verification plan and report (DVP&R), includes a list of verification and validation activities (including design changes, CAE analysis and testing), starting and finishing dates of these activities, in which product development phase the activities will be performed, which item of the whole product is subject of the activity, and what parameters are measured.

4.4 Reasons for testing

Depending on the nature of the product, companies can have different reasons for testing. In the product development process of the case study company, there are several external and internal reasons. The major reasons are described in this section.

4.4.1 External /exogenous

The case study company's testing is largely driven by legislation, market needs and business affordability. In the company, these are termed as "voice of regulation", "voice of customer" and "voice of business".

4.4.1.1 Voice of Regulation (VOR)

A major driver for new product introduction (NPI) is legislation concerning pollutant emissions. The case study company needs to decide on the technologies that it will use to control the emissions. The legislative requirements vary based on the specific market and regulations imposed in that market. These requirements force the company to introduce new technology or technology additions/changes. Legislative requirements provide the guidance and procedures to develop and test a new engine. The company follows related regulation guidance and specification for testing.

A specification is defined as a list of tests, references to analytical procedures, and appropriate acceptance criteria, which are usually numerical limits, ranges, or other criteria for the tests described. The specification establishes the set of criteria to which an engine should conform if it is to be considered acceptable for its intended use.

4.4.1.2 Voice of Customer (VOC)

Engineer 1 mentioned that

> "the company's key strength is its ability to tailor engines precisely to meet customers' requirements, which is why its engines solutions are trusted by more than 1,000 manufacturers".

Although the company has a wide range of customers from different applications, from the customer's perspective, the diesel engine is a mature and stable product. Each version of engine is designed to meet specific customer requirements within the general capability of the engine. For example, for a particular version of engine a customer might specify a power range, load, and pattern of intended use in a machine. There are also non-functional requirements like service, maintenance, safety etc. These functional and non-functional requirements determine how the engine should be developed and tested. The company's engines are used in different sectors which have different environments. These different use scenarios dictate different test specifications.

Market competitiveness also provides an important context for testing. For example, many companies can offer a four-cylinder engine which pushes the limits of performance, reliability and durability and which makes the testing context crucial. Because the engine will be working at the boundaries of its capabilities testing which is closely related to the contexts in which the engine is to work is critical for a successful competitive product.

4.4.1.3 Voice of Business (VOB)

From the business perspective, one of the key decisions that is made during the concept selection phase is the choice of component technology that will enable the product to perform to its specifications at the lowest cost. These decisions determine the types and levels of effort that will be required to devote to develop and test that technology and the engine in which it is to be used.

VOB plays an overall role of selecting and determining different aspects of testing, especially through feasibility analysis of time and cost. Being a subsidiary part and the supplier to its larger parent company, the case study company also needs to follow the overall testing practice as well as guidelines for business processes established by the parent company.

4.4.2 Internal /endogenous

Internal company reasons of testing are mostly related to the company engineers' technical understanding of the product and related technologies. In order to achieve product-based improvements in market penetration and to further the company's technological competencies, the company's R&D team investigate new technologies. The company might also consider a technology that has the potential for improving the level of performance at the same or lower cost as the existing proven technology. At the beginning of a new product development process, testing is driven by twin priorities of "newness" and "problems identified in previous products". Design changes made during the product development process also drive testing. These internal drivers of testing in the case study company are now described under the categories of newness, problems and change.

4.4.2.1 Newness targets

Engineer 2 revealed that in the company

> "We define our staring point as we call newness- simply, what is new about the engine"

Initially newness can arise from requirements of customers and regulations as well as internally from technological innovation within the company. At the beginning of the product development process, newness is related to requirements, technologies and components.

New requirements

New product development (NPD) processes often starts with a list of new requirements. Engineer 2 mentioned that most engines are 18-25% new in terms of how much they vary from previous versions. New requirements arise in three ways from customer, from regulation and from inside the business.

(i) *New customer requirements* are usually carried through from a previous version of the engine with adjustments of the key characteristics, such as "increase power by up to 40% over previous versions, with a corresponding torque increase of up to 60%". The area of application and the environmental conditions are also be new to the company. New customer requirements are mapped across the design specification of the engine and the relative importance of newness of different parts of the engine is weighted relative to the expected application.

(ii) *New regulation requirements* are based on the market where the product will be used. These regulations requirements drive the technologies that the company needs to adopt or introduce, for instance, Diesel Particulate Filter, Cooled Exhaust Gas Recirculation, Oxidation Catalyst and so on. These regulation requirements also include specifications of performance that an engine needs to maintain during its lifetime. These requirements help to define the process of developing and testing the engine.

(iii) *New business requirements* may determine new ways of producing the engine with minimal cost, while satisfying customers and meeting regulations. The company also need to work on upcoming regulations and forecast customer requirements. The company's New Technology Introduction (NTI) programme investigates potential technologies to meet these new business requirements.

New technology

The use of new technologies has been described to be a vital and recurring problem for this company, because the company might have limited knowledge and experience with the new technology. Several situations can occur in New Product Development:

(i) *No new technological requirement*: When the product, the application and the environmental condition are known, there are no new technological uncertainties;

therefore company identifies ways of improving testing processes so that the overall product development time and cost can be reduced.

(ii) *Technological adjustment/ improvement*: The product is well known (proven in use), but the field of application is new. Examples in the company might include using an engine for a mining system, in extreme weather (±40 degree C) or in a desert where sand can be a cause of engine failure. The company focuses on the testing that will prove the product's compatibility to that application and environmental conditions.

(iii) *Technological challenges*: The product is known but has a limited field history, and will be used in a new application and/or in new environmental conditions. Then the product needs to be proven in both its technological capability and application compatibility.

(iv) *New Technology Introduction*: The team, in particular, may consider a prospective technology that is not yet fully proven but offers the potential of a superior level of performance at the same or lower cost than the best existing proven technology. In this case, testing needs to demonstrate the products technological capability as well as being proven for intended applications.

New components

The company might wish to procure a new component, like, a filter, which will be customised to their needs from a known supplier. A new component or part needs to be proven to be compatible not only with the rest of the product but also compatible with the operational and environmental conditions. The process of testing and verifying a new component from a new supplier can be significantly different from a proven component.

4.4.2.2 Problem in the previous product

A problem can be identified after the development phase during manufacturing or in use. Problem in the previous product means that it did not meet a customer target or functional requirement. As Engineer 3 mentioned, the company focuses on functional requirements,

> "when it has not met what we felt the necessary target was to meet the functional requirements".

Problem in manufacturing

Although, the company understands its own manufacturing process and considers in great detail the manufacturing variability during design and development, there are still many problems can arise during manufacturing. Material properties vary batch to batch, assembly process and sequence change or tooling wears over time and use. At the start of the design of a component, a full understanding of manufacturing plan and the processes that is going to be used is critical. In this way the company can plan and closely control manufacturing variability. Manufacturing problems drive different testing methods and acceptability criteria back into engineering design.

Problem in use

Any failure occurring in the field is considered as a high risk. Issues identified in use significantly drive next generation product development and testing procedures. The company continuously monitors and captures a product's performance and durability when engines are used in a field. For a new product development, the company uses information from the 'use in the field' (how product is performing in use) and the 'use of the customer' (how customer is using the product) to judge what are the likely occurrences of any potential failure.

4.4.2.3 Emerging changes during product development

Changes occur in response to problems arising during the development process. These can be design changes, requirement changes or part/component changes. Changes have significant effects on how the product is designed, developed and tested. Engineer 1 mentioned

> "The design change has a propagated effect which eradicates some of the testing, introduces more testing, questions whether we have tested in a right way that is the thing that destabilised the process"

Change in design

Initial design can consist of problems and mismatches with requirements; therefore testing is required to identify those problems and/or to corroborate the design. Any problem in design forces design changes and further testing. Problems identified in testing tell how much to change a design, which then forces the company to investigate how to plan and perform the further tests to validate these changes.

Changes in customer requirements

Initially, a customer requirement can be unstructured or ill-structured. A customer might not be completely clear about their expectations, or sometimes the company cannot fully understand a customer's requests. As the development time of engine last spans 3-4 years, the company may find customers want to change a requirement, ask for something new or even want to remove some requirements. There can be changes of requirements after contracts have been signed and while the product is being designed.

		Understanding of requirement	
		Structured	Unstructured
Customer request	Solicited	No change or little change	Moderate change
	Unsolicited	Can Change/ need to eliminate	Can completely Change/need to eliminate

Figure 4.18 Changes in customer requirements

Changes of customer requirements can be modelled using two factors: 'understanding of requirement' and 'formal contract/agreement has signed' (shown in Figure 4.18). 'Understanding of requirements' can be structured or unstructured. The understanding of a requirement is structured when the requirement is clearly defined by the customer and the company has a clear understanding of the customer's expectations. An unstructured/ill-structured understanding of requirement means either a customer is not sure about his/her wants or company is not completely clear about the customer's expectations.

Changes in requirements cause design changes; therefore, the associated testing process also changes. If the requirement changes after a formal contract or agreement has been signed, the consequences are far more complicated than before signing the agreements. This is because, when a contract is signed, the company formally starts the subsequent process of designing and procuring items, especially of prototype parts and components for testing.

Changes in parts

During procurement for a prototype component, a supplier can inform the company that the component doesn't work to its specification. A company engineer then has to change the design. An example is a controller that is going to be used in the engine turns out not to have the required sensitivity which may lead to more variable behaviour of the design. This results in changing the types of testing and/or the details of the test.

4.4.3 Discussion of reasons for testing

The priorities among several reasons for testing may vary from company to company. In the case study company, regulation enforced tests are considered to have the highest priority and are mandatory for product acceptance. The company follows the regulation guidance and needs to provide the justification for the selection of test procedures and acceptance criteria for new engine development. Several other factors affect the testing decisions. For instance, engineers might consider that it is necessary to perform a series of tests; Test1, Test2 and Test3, in order to prove the technical capability of the product, whereas Voice of Customer and Voice of Regulation might enforce Test4 as mandatory. Voice of Business is the factor that finally decides the overall yes/no decision of selecting of tests. By analysing the "needs" of engineers and "musts" from regulations and customers, a validation manager might decide to select Test1, Test3 and Test4 based on the feasibility and business affordability. Engineers might still decide to perform Test2 in a different mode, such as modelling and simulation.

Overall in the company is guided by the priority of risk reduction. As Engineer 3 remarked;

"Product development process is a risk mitigation exercise"

The design process can be viewed as a series of tasks and decisions for reducing uncertainty; uncertainty is termed "risk" by the engineers in the company. A product development project has high risk when the technology is not well understood or a product's capability ability is not known. At the beginning of a project, potential applications of the product and the environmental conditions under which it will be used are not fully defined. Further, the process for implementing the project's goal may not be clear at the start. The company assesses risk by asking three questions: "what is new", "what has failed previously" and "what needs to change". As expressed by Engineer 1,

"For the core engine, we start where we left off with the previous engine and first of all we define where we're not very good"

"Newness" is considered as the highest risk since the product and process are not clearly known. At the beginning of the product development process, risk arises from the multiple alternatives to choose from. Technology selection can be particularly challenging because it is difficult to predict with any certainty how the final product will perform or which alternative to take to deliver balanced costs and benefits.

The company does not classify, initially, any qualitative difference between a new functional requirement and a functional requirement with previous failure history. Both of the scenarios will get the same attention for risk minimisation. To reduce the cost and risk associated with a new product, the company limits the "newness target" in the product. The "newness target" in terms of new components, technology or reuse in different contexts, introduces risk into the system and can prove to be challenging for the company. Risk associated with "change" in a part or a component is calculated in terms of a component's complexity, interaction with other components of the system and the company's familiarity with the item.

Analysis of changes is critical as some changes propagate. This may have serious consequences resulting in the need for radical changes in design and testing. Development testing is performed to eliminate these uncertainties and to ensure that designs meet requirements for performance, safety, durability, reliability, as well as regulatory and statutory aspects.

4.5 Objectives of testing

Each of the internal and external reasons has several objectives behind recommending or enforcing any particular test. External reasons enforce tests mainly to verify and validate engines against customer requirements and related legislations. Internal engineering reasons include, performing tests to learn about the effect of implemented changes in engine or to investigate how technology can meet the newness target. Also, at different stages of the product development process, there are different objectives of performing tests. The following objectives of testing were identified in the case study company, where testing for learning was repeatedly highlighted as a key objective.

4.5.1 Testing for learning

Since new product introduction (NPI) presents many uncertainties because engineers possess insufficient knowledge or information about the product and the process, they may be uncertain about a product's performance and behaviour. Testing is required to reduce uncertainty in knowledge and information about a product's functionality and performance, the behaviour in a certain conditions or interference within the system. A lot of testing, at the concept development and selection phase, has an overall learning objective aimed at reducing or eliminating epistemic and aleatory uncertainties.

4.5.1.1 Epistemic uncertainty reduction

According to McManus and Hastings (2006), lack of knowledge or 'epistemic uncertainty' refers to "facts that are not known, or are known only imprecisely, that are needed to complete the system architecture in a rational way". Testing can reduce these underlying knowledge uncertainties by creating, collecting and confirming the information about a product in several ways including:

(i) *Providing knowledge*: Engineers might simply not know how a product will perform in a certain operating conditions- they need to do functional testing to identify the actual behaviour of the product. Testing can provide knowledge about the product's actual behaviour, including interferences between modes of behaviour as well as unforeseen emergent behaviour.

(ii) *Defining knowledge:* During the design analysis of an engine engineers may only be able to predict a parameter's value in a range; for instance, engine power might be between 37-42 kW or fatigue properties may be approximately assessed. By the measurements from a performance test, actual values of the engine power range can be identified or precise fatigue properties of a component's material can be known through a mechanical test.

(iii) *Providing trust in knowledge:* Engineers can have limited confidence in their design decisions, especially when the product and process are new or have changed significantly from the past. They might well know the technology but not sure how the product will behave in a new or changed environment. Therefore, testing can reduce these uncertainties and can increase trust. As Engineer 4 mentioned, "testing builds confidence".

4.5.1.2 Aleatory variation reduction

Aleatory variations are sometimes termed as objective or stochastic uncertainties (Wynn et al. 2011). These are inherent variations associated with a physical system or environment; such as dimensional variation in machined components. Often the manufacturing of the prototype product and production intended product are not completely uniform, which introduces aleatory variations i.e. uncertainties. Aleatory uncertainties are random and can be irreducible (Aven & Zio 2011) but testing measures these variations and can identify how much the actual product varies from intended design.

4.5.1.3 Profiles of learning

Learning results from conducting a range of tests, each with different characteristics. Durability and reliability tests usually take a significant amount of time because these tests are designed to mimic the products' life time behaviours. Functional performance measurement tests can be carried out relatively quickly with higher frequencies. For example, during the "Lubricating Oil Pump Performance" tests, the speed of the pump and oil flow are measured by varying the delivery pressure and oil temperature. These readings are taken over a 20 second interval. Multiple measurements are taken for different oils, in different temperatures, to establish the pressure/flow characteristics of the lubricating oil pump. To establish the capability of the lubricating oil pump using these performance tests can take approximately 250 hours, when all the variations and repetitions are taken into account. On the other hand, a "Lubricating Oil Pump Endurance Test" ascertains the durability of a lubricating oil pump; usually runs for a total of 1000 hours at a set oil temperature and at an equivalent to engine rated speed.

The likelihood of finding faults through a test is a bimodal distribution with many faults showing up early in the test and others towards the end. However, the state of the system or its components can only be assessed once a test is completed. While a test might not fail as such, the state of particular components might not be acceptable and require redesign. For example a piston might show signs of excessive wear when inspected at the conclusion of the test.

The profile of information acquisition varies in different tests. In some tests, most of the changes happen at the beginning and information evolves to a final value rapidly. In others, information evolves slowly at the beginning and increases most towards the end. Usually, in a functional performance measurement test, the final value of the measured parameters can be attained quickly, so functional performance measurement tests can be considered as fast evolution tests. But in mechanical durability or reliability tests are usually lengthy, and information evolves either gradually throughout the testing process, perhaps increasing slowly towards the end. For example, a "Deterioration Factor" test in the case study company, checks how performance and emissions change over time. This Deterioration factor test is usually run for 1000 hours. Engineers learn the behaviour of the product gradually as the process progresses and information evolves at a regular and continuous pace. On the other hand, in a "Gross Thermal Cycling" test that runs for 1000 hours; most of the learning is attained towards the end of the test. However, in both mechanical and functional tests, if a component or system fails at the start of a test, engineers can immediately learn something from this 'infant mortality' (see Figure 2.3), which is passed on to the downstream activities. In general the company tends to acquire most information from tests at the beginning and end of the test (Figure 4.19 illustrates such a profile).

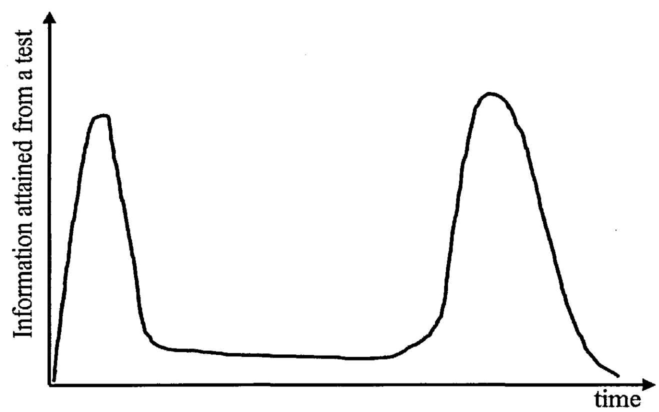


Figure 4.19 Illustration of information acquired from a test over time

4.5.2 Testing for Demonstration

When the case study company engineers need to generate and evaluate available alternatives, testing is a tool for demonstrating the feasibility of the concept and to prove 'the concept' to be right one to select. This is especially the case when the company has limited knowledge about a new technology, which can have potential benefit over a proven technology. Engineers need to demonstrate the technology to the business managers by citing evidence from testing to assure feasibility.

4.5.3 Testing for Verification

Design verification aims to ensure that design outputs meet the given requirements under different use conditions. The company's verification process ensures that the product has been built according to the requirements and design specifications. The verification is accomplished by a combination of test, measurement, analysis, review of design and inspection. Testing for design verification objectively provides evidence that specified requirements have been fulfilled.

4.5.4 Testing for Validation

Validation is a process to proving that resulting product is capable of fulfilling the requirements for the specified application or intended use. Testing is an essential technique for product validation, as Engineer 2, a validation manager, comments:

"Testing reveals the truth".

The company's product validation process involves testing the product against customer requirements and specifications, e.g. under a range of potential use. The validation process also confirms that a product is manufactured with the appropriate production processes.

4.5.5 *Testing for Certification*

Global emission regulations for diesel engine manufacturers provide requirements for the testing and evaluation of new components and new engine designs. It is an imperative for certification that the company follows the standard regulations during product development in terms of how a product needs to build and tested during validation for certification. For instance, to meet the in-use compliance, the company needs to demonstrate that the engine will meet specific levels of particulate emissions that will be detected and measured at the end of the useful life of the product.

4.5.6 *Mapping between reasons and objectives for testing*

Each of the tests performed in the case study company has a reason and an objective or purpose. From the evidence of the Case Study, Figure 4.20 shows how reasons can be mapped against objectives.

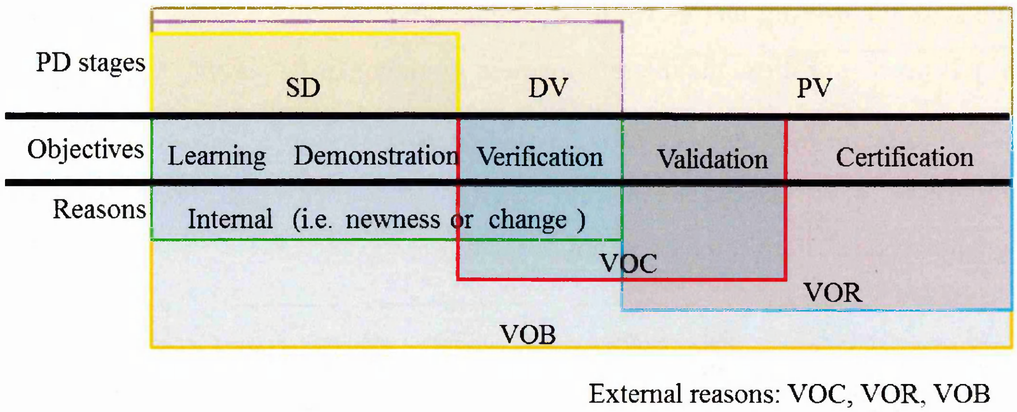


Figure 4.20 Testing objectives are mapped to testing reasons at different product development stages

Internal tests are for learning, demonstration and design verification. Externally, customers are demanding tests for verification and validation. Voice of regulation (VOR) mainly monitors how customer requirements and environmental factors are respected during design. These tests are required for certification. Voice of business (VOB) monitors the optimal balance between internal and external demands as well as contribution to the business in terms of profit.

4.6 Comparison with another company

To broaden the understanding of this company specific analysis, it was decided to compare the findings with another company. The intention was to answer two specific questions; first, when does physical testing happens within the product development process and second, what is the role of Computer Aided Engineering (CAE). To avoid confusion, the original case study company is called as Company A, and this second company is called Company B in the following sections. The study in Company B was based upon two semi-structured interviews with (a) the Test and Validation leader and (b) a mathematical modelling and simulation engineer. In addition there were several informal discussions with the mathematical modelling and simulation engineer.

Company B was established in 2011, but works under the brand of a parent corporate company, designs and manufactures counterbalanced forklift trucks. These forklift trucks are designed to meet the needs of light to medium duty operating environments. Compared to engines, a forklift truck is relatively simple product with straightforward functionality. For Company B, the regulations related to safe use and operation of the fork trucks are the driving factors for new product development. As with Company A, the Company B needs to follow the overall business guidance of its parent.

Product development process (PDP) and testing

Company B has a six stage-gate process structure shown in Figure 4.21. It is called a 'review gate process' within the company.

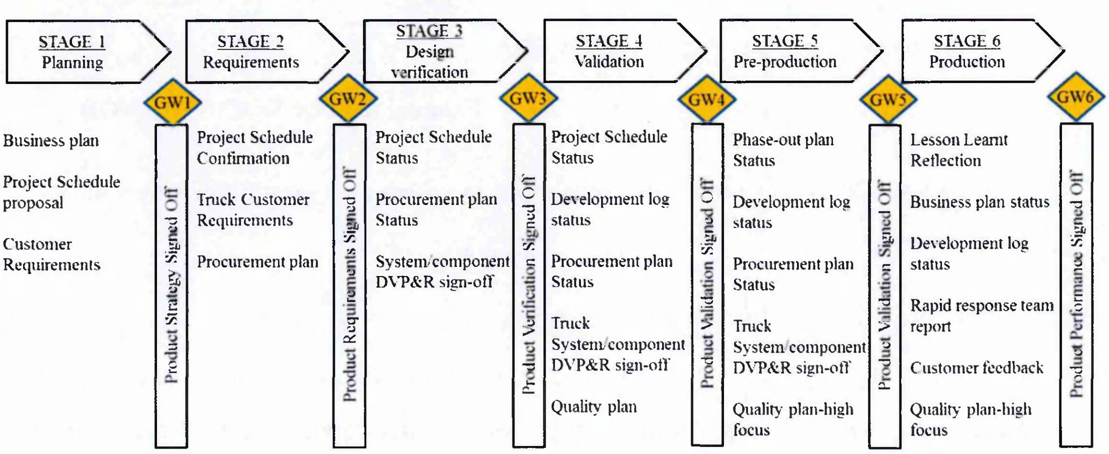


Figure 4.21 Product development process structure Company B

Figure 4.22 shows how testing is performed throughout the product development process in Company B. The initial concept design is analysed through CAE and simulation and modelling. In the 'Design Verification' phase, a 'MULE' truck is produced with a combination of new and old components and physically tested to verify the design.

'Prototype A' is built and tested during stage 3 to stage 4 to test performance mainly. Finally, 'Prototype B', which is production tooled, is tested in stage 4 to stage 5 for product validation and certification. Therefore, Company B also starts physical testing at the early stages of PD and there are at least three iterations of prototype testing during the product development. This PDP has a similar structure to Company A (shown in Figure 4.5) and shows that testing is closely intertwined with design in each stages of PD.

Company B extensively uses CAE analysis during the concept development phases, such as structural analysis, hydraulic modelling and simulation, before committing to prototype building. These analyses are particularly used to explore the design opportunities and for concept selection. However, at a later stage, during design verification and product validation, Company B largely depends on physical testing.

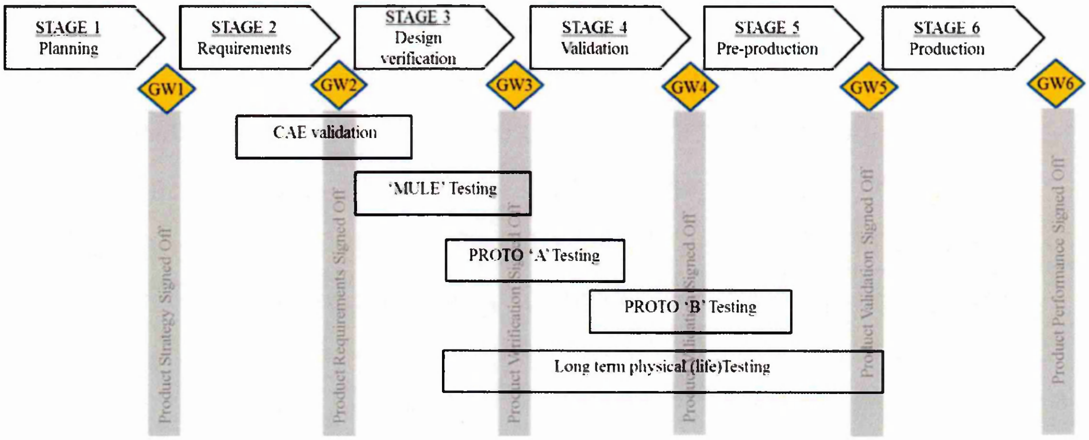


Figure 4.22 Testing in the PDP process of Company B

CAE helps to perform the testing because the CAE simulations and models give the pointers to look for. Several are based on some form of finite element analysis, particularly stress analysis. As the modelling and simulation engineer commented:

> "The role of validation is to validate the FEA results, there might be 20/30 variations of one product, we can't build and test all of those, we may build three or four variants, if we can validate CAE or FEA against those physical trucks we have built, then we can sign off the entire range"

Increasing the development and use of these models and simulations means that the company acquires more confidence in FEA for building the next new product. It may be possible to reduce to the number of physical products that are built and tested while signing off more designs using models and simulations.

4.7 Summary of empirical study in the case study company

This empirical study highlights several key points:

1. The testing process is an integral part of the PD process and closely intertwined with design.
2. To integrate testing throughout the PD process upfront analysis of QFD and FMEA can play a vital role in linking design considerations with the testing plan.
3. Different reasons affect the process differently. For example, external (sometimes referred to as exogenous) forces, like legislation, make some tests mandatory; and thereby determine or shift the testing practice in the company. Whereas, mitigating risks associated with new technology is an internal reason for testing and is to some extent under company control. The ways that these internal (or endogenous) reasons affect the process can be balanced against competing claims on resources to assure viable and competitive products.
4. The observations from Companies A and B emphasise the importance of CAE in current improvements to their product development processes. This prompts a reappraisal of the nature of CAEs in order to get a clearer view of how it can be used in increasing the overall effectiveness of testing as well as the links between testing and design activities.

Chapter 5 Framework for test planning in product development

In the previous section, the reasons and objectives for testing were set out and classified, based on the examination of the case study company. These answer the question of why a test needs to be performed. This chapter enumerates four entities, which have emerged from the case study as the main characteristics of a test about which decisions need to be made in planning the testing activities. These relate to the objects being tested, their properties, the mode and location of testing. During testing planning the company makes decisions about these in specifying a test:

- what is the *object* under a test (for example, a component or a sub-system),
- what is the *property* that is measured in that test (for example, durability or performance),
- what is the *mode* of that testing (for example, physical testing, CAE analysis or simulation) and
- what is the *location* of the test (for instance, in-house, in the supply chain or out-sourcing).

In this chapter these factors in test planning are analysed and a framework is presented for planning testing by establishing the ways that these entities, namely, object, properties, mode and location, are considered in the test planning process. The framework is developed from the analysis of the data from the case study and contributes to understanding the processes of testing. However, most significantly it provides the basis for proposals in the thesis for ensuring that the substantial resources required for testing are deployed in ways that are most effective in delivering timely and relevant information to design and product development.

5.1 Entities characterising tests

The first section present details of the objects, properties, modes and locations from the case study company, which then forms the basis for constructing a framework for test planning.

5.1.1 Objects

An engine is a complex product consisting hundreds of parts, components and peripherals. The case study company's testing occurs at component, sub-system and system levels. Testing the critical parts and components is vital to identify their functional ability, whereas sub-system level or system level engine testing are essential for integrity and durability of these components.

A major decision that the company needs to take is around the balance between component, subsystem level and system level testing. Component and subsystem level testing allows parallel testing thus swift fault recognition. However, while a module/sub-system might meet its own specifications, unpredictable interference and/or malfunction may appear when the whole system is assembled. This decision on the balance among component, subsystem and system level testing has an influence on what to test, as well as on when to test.

5.1.1.1 Components

A component can be tested at different levels: at a component level, a test is performed on the component; at a subsystem level, a component is embedded into a subsystem and tested but the focus of the test remains on the component. Similarly, at a system level, a component is tested by putting it into an engine under test.

For a new component, the capability to conduct performance and functional assessment of engine components is assessed by component tests and experiments. Engineer 4 mentioned that:

> "component testing can often be more cost and time effective than one engine testing. In some cases, it is the only way in which component properties can be reliably determined".

The company's Mechanical Engineering Division consists of four main teams corresponding to the different areas of the product: Air & cylinder, Rotating Components, Cooling & Lubricants and Majors & Castings Groups. The teams work together during product development and testing. Each component of the engine is assigned to one of these teams. four categories within these divisions.

5.1.1.2 Modules/subsystems

An engine will have several sub-systems, for example, the fuel system, the cooling system, the air intake system and the exhaust after treatment systems. Performance of each sub-system is ensured through functional and mechanical testing at several stages of product development. This usually happens at engine level and machine level testing.

5.1.1.3 Product/systems

The Company takes a systems engineering approach during product development and testing of the engine. The engine is considered as a system consisting of hundreds of components and sub-systems which are tested to ensure the overall product's performance and durability. In the product development process most of the design verification and validation testing is engine level. Engines also experience machine level testing at each of the product development stages.

5.1.1.4 Peripherals

There are several accessories attached to an engine. For example oil pump and fuel pump, go through several testing stages. Initial testing of these peripherals is usually performed by the supplier, according to the specifications provided by the company. However, the performance of these peripherals needs to be assured through module, system and machine level testing, because the interference with other modules/subsystems or components can produce unexpected or new behaviour of these peripherals.

5.1.2 Properties

Initial "Voice of Customer' (VOC) and 'Voice of Regulation' (VOR) requirements are translated into an engine's technical requirements and those for its components. Testing measures the acceptability of the product against those requirements. Engineers identify and decide the parameters that are being tested in a particular test procedure and the specific measurements they need to take. The company mainly categorises testing with respect to the two key product properties: functional and mechanical. The terms "functional" or "performance" can be different in other applications. This thesis employs these terms as they were used in the company and these are described below.

5.1.2.1 Functional properties

Functional tests measure, ensure and corroborate that engine under development will deliver its required performance, while ensuring that the engine meets applicable legislation. A company team called 'performance and emission' (P&E) are responsible for ensuring the product's performance and emission through measurement and test.

5.1.2.2 Performance measurement

Performance testing measures engines properties. For example, power and fuel consumption of an engine may be measured given a regulated fuel and air intake into the engine cylinder under steady state conditions of constant speed and load. However, many customers require their engines to run transiently from one speed to another, and across changing load conditions. Different applications require different torques at different speeds. Different weather and altitude also dictate the performance measurements that are required. For example, in cold conditions engine's performance changes because engine friction increases as oil viscosity increases. At a high altitude the density of air reduces, and less oxygen in the cylinder means less fuel can be burnt, resulting in less power being produced. Therefore, engineers calibrate engines such that at higher altitudes they will be able to produce the required power and operate within the design limits of the engine. Performance measurements ensure an engine's functional performance in wide range of applications.

5.1.2.3 Regulation compliance

While ensuring the performance, engines need to satisfy legislative conditions. While measuring the performance parameters, engineers measure the chemical constitution of the exhaust gases. Performance and emission testing are performed concurrently to ensure the optimal balance between transient response and emission, for instance. Because legislative requirements are so critical, the company invests in the latest diesel engine calibration equipment, which improves their overall testing capability.

Emission control

The performance and emissions (P&E) group manage performance measurement and emission control related to NOx (Oxide of Nitrogen) and particulates emissions. In order to meet desired NOx and particulates emissions, engineers calibrate and change control variables. These variables include the number of injection shots, injection pressure, shot duration and other parameters. The performance of the engine is optimised using statistical modelling of engine input variables and/or test components and modules on the engine on a test bed.

Noise vibration and harshness (NVH)

To meet the legislative requirements on reducing noise emissions, the company has a team called the noise vibration and harshness (NVH) department. The NVH team carries out noise tests. Together with P&E team, this NVH team investigates the different technologies, for example, fuel injection developments, and ensures that relevant regulations are met through testing.

Vibration can often be a cause of noise. The NVH team's work is involved in supporting the development of new engines, ensuring that vibration levels are within acceptable limits to prevent component failure. Through testing, this team identifies noisy components and subsystem modules and redesigns them to reduce overall engine noise.

Different customers will feel the effect of noise and vibration differently. This perception is termed "harshness" in the company. Engineers want to reduce these effects by making the engine and machine sound better to the customer. Through testing and analysis, this team predicts the effect of different noise treatments on the overall noise emitted by the engine.

5.1.2.4 Mechanical properties

Mechanical testing is mainly performed to ensure the reliability and durability of the product. These tests are conducted on components, sub system modules, on an engine and/or on an engine mounted in a machine, testing over long periods of time. The company's mechanical tests are conducted at much harsher conditions than the engine will see in the application. Both reliability and durability testing depend on the product's application and environmental conditions in which the engine will operate.

Reliability

Reliability tests ensure engine's ability to perform the specified operation without failing over a period of time. Engines are required to perform reliably over a 10000 hours of operation life. The company must ensure that all specified components are adequate to function through the intended product life. Engineers conduct reliability tests to understand and find any variability of engines from specifications. These results are then statistically analysed and extrapolated within the variability of design specification and manufacturing capability to obtain estimates of life expectancy or in service under normal conditions. Testing during development verifies that the product can be deployed with a low failure rate. With a limited time frame to develop the product, the company accelerate the testing methods to demonstrate that the product will achieve the desired life, although it is expected that the stresses will cause the product to degrade or change their performance over time. However, it is critical to know how long until the probability of failure starts increasing and product reach to a wear out phase. The case study company uses the accelerated testing methods. Most of the accelerated testing is to verify that the product will perform reliably during the useful life, until they start to wear out. Very exceptionally and highly undesirably, some products/components fail at the start of a test due to infant mortality. During the tests, the performances of a product are measured, and how the performances change over time, are monitored.

Durability

Durability tests assure that, as Engineer 1 commented:

> " the product will work- given proper maintenance, for a given duration"

These durability tests are conducted in peak harshness and tougher condition for a reasonably short period of time, called accelerated tests, forcing components or engine to fail/pass. For example, a gross thermal test procedure specifies the test cycle for determining the thermal fatigue resistance of core engine components. The test cycle does this by subjecting the engine to a controlled, rapid coolant temperature change cycle. It is intended that when an engine is run for extended periods to the test cycle given in this procedure, it will simulate the conditions met in service over the full engine life.

5.1.3 Modes

Tests can be performed in different modes or in a combination of modes. There are mainly two modes of testing recognised in the company, (1) CAE analysis through modelling and simulation to predict the properties and behaviour of an object and (2) physical tests to evaluate an object's actual behaviour. CAE analysis can be used for learning and demonstration and for design verification to some extent, whereas physical tests are required for product validation and certification.

5.1.3.1 CAE Analysis

CAE is playing a significant role in the case study company's design process, as Engineer 7 comments:

"CAE is becoming increasingly important to the companies to minimize the effort and expense involved in product development".

The company has long experience of using CAE and realises the benefit of using these analyse. As Engineer 2 mentioned:

" the design team can iterate the design process to develop a product that better meets cost, performance, and other constraints".

At the early stages of a product development programme, the CAE analyses are used to investigate different trade-offs, usually 1-D analysis (a 1-D simulation model is a mathematical representation of a system and its dynamic behaviour). These models allow the engineers to simulate and understand the interaction between components, for example how an engine performs in a context, when given a load requirement for speed and acceleration. From these analyses "design briefs" are created. According to these "design briefs" individual components are designed and component level CAE analyses are performed to structural capabilities, for example, of meeting those "design briefs". Once the component level design and analyses are finished, there are system level virtual assessments of the whole engine where the outputs from the individual component level are considered to create and/or update the system level CAE models. These system levels CAE analyses prove that the initial concept is going to work.

An example of this process is a thermal-mechanical analysis of a combustion model; the initial trade-off simulations are performed to understand the combustion requirements that are needed to produce required power and torque. These CAE analyses answer questions such as: is this engine going to be the right architecture, how many cylinders are there, sizes of the bores, how much fuel can be put through the system. Also, the simulation of the model indicates what is the thermal load, or thermal input to the piston or the cylinder head. These analyses give design limits for pistons or cylinder heads. Detailed design activities are completed based on these initial design limits. Component level CAE analyses are performed to prove that the piston is designed according to those design limits. Once the detailed design and analyses of all components are available analysts perform full thermal mechanical analysis again, because, in their initial analyses all the detailed information about the components or exact measurements were not available. In this iteration of the full thermal mechanical analysis, they start to analyse the interactions and effects in detail, for example, is this prediction of how much heat is going to the piston, or what is the effect of that going to be. So when the actual validation tests are performed, usually engines are on test-beds, the engineers will be able to measure all the boundary conditions that they have used in CAE models, and confirm the correctness of their assumptions.

From the above scenario, it appears that there are several levels of CAE analysis that happen in a company: one is system level, which is model based or 1-D analysis to understand the system interactions and working principles. Another is component level, which are detailed analyses performed after design works start and often in parallel to design. These detailed CAE analyses typically fall into three main areas: structural analysis, mechanism or dynamic analysis and thermo-fluid-flow dynamics. They result in the determination of parameters like material properties, geometric idealization, and physics, which help to define the scope of the design activity. These detailed analyses also identity an initial set of boundary conditions and operating conditions, which are compared with design briefs to refine the product requirements and improve through design iteration. In summary the case study company uses CAE analysis to:

a) support earlier design decisions,

b) explore the design opportunities by varying the parameters and

c) meet the design parameters.

5.1.3.2 Virtual testing

A further level of CAE analysis and simulation are performed to identify the behaviour of the systems/components in response to the specific environmental conditions. The later types of CAE analyses are used to narrow down the boundary conditions and provide specific information to the physical test engineers. For example, in a performance test, simulation can predict when to measure a value or in what conditions, and predicts the value that will be measured in a physical test. So test engineers can spend less time on the physical test accurately knowing what is expected. If the expected values do not correspond to test measurements engineers can assume that either the analytical method applied for virtual testing is not accurate or there can be mismatches between the test settings and the CAE. The case study company's physical testing depends on virtual testing before components, modules or systems go to actual physical testing. Simulation is a space where the physical testing is optimised and virtual testing helps to focus on the conditions that are need to be physically tested. Engineer 1 mentioned that

> "twenty hours of focused testing is better or equivalent then thousands hours of non-focused testing".

The company may choose to carry out a physical test for a baseline product while using simulations for multiple variations developed for specific use cases. The baseline product is the standard product without adjustments for specific needs and its testing is described by Engineer 1:

> "the baseline product definition is physically tested and that information is fairly adequate for simulation to run for multiple variables for longer time to find the optimum setup. Then a physical test is required to validate the product as well as simulated results".

The validation manager (Engineer 2) mentioned that by feeding the physical testing results into simulations, a virtual test can be used to design an optimum system with the desired static and dynamic characteristics. These types of analysis simulate object properties, which are tested in physical tests, as well as drive testing process, through identifying those product and component characteristics that require testing. Subsequent physical testing is used to validate simulations.

When designing a test setup for a mechanical product, test engineers use best practice, experience, and expertise to determine methods and objectives. However, in recent years, the significant improvements in CAE mean that CAE analysis can now provide load cycle, loading locations, sensors locations, sensor and system calibrations, for example, to the test engineers. In summary, CAE analysis is also applied to understand the behaviour by varying the environmental and operating conditions as well as to set-up physical test conditions, input parameters and sensors locations.

As the company's engines are used in many applications, operating conditions also vary across customers and conditions of use, the company can only afford to test a limited number of variations, which are usually the worse-case scenarios. But the other variations can be tested virtually through CAE simulations. Consider the example of a proven technology deployed in a new context (for example, in different use conditions and environment. It needs to be tested in these new scenarios. Virtual testing can achieve results across the range of scenarios.

In summary the case study company uses virtual testing to:

a) understand the behaviour by varying the environmental and operating conditions,

b) to set-up physical test conditions, input parameters and sensors locations and

c) to assist physical testing.

To recap, in the case study company, design processes begin with mathematical models of the target performance which are increasingly expanded as more and more factors are included in the models as part of a requirements cascade (Wyatt et al. 2009). The Computer Aided Engineering (CAE) team runs the initial performance models. The geometry of the components is designed around these performance models by component teams and later handed back to the CAE team for validating through Computer Aided Engineering (CAE) tools like FEA and CFD. The models are further refined to simulate load cycles, loading locations, sensors locations, sensor and system calibrations etc. as an input to the test engineers. As illustrated in Figure 5.1, this is a continuous development and use of computer models, but with distinctly different purposes.

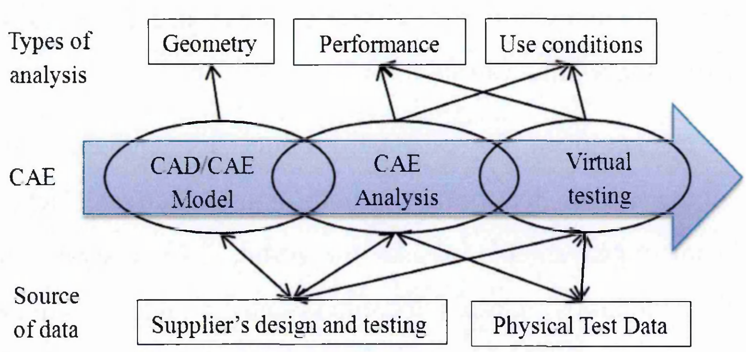


Figure 5.1 The progress of CAE and interactions between types of analysis and the types of data from design and test activities.

As shown in Figure 5.1, there is no indication of when a CAE analysis becomes a virtual testing. The transition from 'CAE analysis' to 'virtual testing' is rather fluid and is not a straightforward linear process. Also, the terms "CAE analysis" and "virtual testing" are used interchangeably in the literature. But this research has identified that there is a subtle difference between these terms, which can only be captured when they are used in a context. In the following sections these terms are analysed using the examples from the case-study company in terms of:

i) What is the process involved in these CAE analysis and virtual testing, and who are involved

ii) when the company uses these in the product development process, The aim is to define virtual testing. A better understanding of the distinctions between CAE analysis and virtual testing help in planning these activities (is discussed in details in Chapter 8).

i) What is the process involved in these and who are involved?

A typical CAE process involves pre-processing, solving, and post-processing steps. In the pre-processing phase, engineers model the geometry and the physical properties of the design, as well as the environment in the form of applied loads or constraints. Next, the model is solved using an appropriate mathematical formulation of the underlying physics. In the post-processing phase, the results are reviewed. This is quite standard practice in many companies (see for example Siemens description of their product lifecycle management (PLM) (Siemens 2014)).

When a model of a part or an assembly is created, the designer-engineer defines a set of boundary and operating conditions, and then typically performs some form of finite element analysis (FEA) to identify the behaviour of the part in response to those conditions. For example, in a static analysis, a given force could be applied and the resulting stresses identified; in a thermal analysis, a source of heat at a given location produces a distribution of temperatures across the part or assembly; and when fluid mechanics are relevant, an initial uniform flow can be influenced by both flow and thermal factors. The results of the analysis might be presented graphically showing expected velocities, temperatures, and pressures.

The analysis of virtual testing processes identifies that both CAE analysis and virtual testing processes comprise of similar activities. The Virtual process also includes these three steps of pre-process, solving and post process. Therefore these analyses don't highlight significant differences between the CAE and virtual testing processes. This is possibly the reason why these terms are often used interchangeably.

Considering the processes of CAE and Virtual Testing in the case study company, initially design engineers and CAE team are involved together in the specific CAE analysis. Design analysis and testing groups work together in a close-loop, facilitated by CAE group. This research also identifies that for both CAE and Virtual Testing, similar tools and software are used. The same engineers execute both CAE and Virtual Testing. They are expert in particular software and forms of analysis, such as, structural analysis, thermal mechanics, fluid dynamics or kinematics and dynamic analysis of mechanisms. There is a close connection between the engineers' expertise and tools, which means that there is a significant common ground between the CAE analysis and Virtual Testing activities.

ii) When does the company use these in the product development process?

As shown in Figure 5.1, CAE analysis is an evolving process and has different applications as time progresses. This research identified that the design objectives, related conditions and depth of these analyses, change over time depending on which engineers or product leaders are defining the tasks or using the results.

Some CAE analyses (1-D modelling and simulation) start even before design itself starts and help to create the design briefs based on requirements. Further, developments of these CAE analyses are performed almost in parallel and iterative way with design to define the scope of the design activity; finally the advanced types of CAEs (which are referred as virtual testing) are performed once the initial design is completed and design data and information are released to suppliers for procurement. These later type examines whether a design meets the specifications and requirements and serves the same purpose as the physical testing of that phase of design or complements and assists physical testing.

At the initial stage, the objectives of CAE analysis are to study the design, based on current understanding of product specifications and design briefs. This initial design and analysis initiates the procurement process of each stages of product development (see Figure 4.6 company's PD diagram). Critical aspects of design are often known later through risk analysis in FMEAs. A supplier's design analysis and testing data also provides information at this stage. Therefore, initial assumptions are frequently changed and CAE models and computational solvers are changed accordingly. Therefore further CAE analysis is performed as the detailed information about the individual components is required or new critical areas are identified. In the later stages of the product development process, the verification and validation team drive the CAE analysis.

Definition of virtual testing

From the analysis above, a phase of the process 'design release to supplier' was identified as a critical phase that helps to define the virtual testing. After design release CAE analysis effectively becomes virtual testing. For the purpose of this research, this is a working definition of virtual testing. This is discussed below in relation to other definitions which are found in academic literature.

Helmreich & Reinwardt (1996) defines

> *"virtual test, i.e. developing, debugging and verifying a test program in a simulation environment rather than on a physical tester".*

Lu et al. (2009) defines virtual testing as

> *"developing and debugging a test with simulation models of the device under test (DUT)".*

Jones & Simon (2002) mentions that

> *"simulation which is performed after a design concept is established leans toward the definition of virtual testing, meaning that in-lieu of the simulation; it would be possible to build a prototype of the design and test it".*

A broader view of virtual testing was presented in (Huizinga et al. 2002) who identified that

> *"closer co-operation between designer, analyst and tester is necessary, and right from the start of the engineering phase the reality has to be simulated by performing complex calculations and analyses using advanced hardware and software. This is referred to as 'virtual testing'".*

In the case study company definitions of virtual testing included a description by Engineer 1

> "virtual testing is all about simulating the test conditions, which is the history of knowing that the product worked. Whereas CAE is all about trying to prove that the product will meet the requirements that the customers, and legislations are asking".

For this research, as a working definition, virtual testing is defined as

> *"an activity of testing in which computer simulations are used under specified test conditions, on the computer model of a physical artefact instead of a physical artefact and virtual testing is performed after a design concept is released".*

The aim of the virtual testing is to evaluate of some aspects of the physical artefact.

5.1.3.3 Physical testing

Physical testing is an activity in which an example of a system, module or component is tested under specified conditions. The results are observed, recorded, and an evaluation is made of some aspect of the system or component. Frequently, engineers in the case study company mention that 'testing builds confidence' or as the validation manger (Engineer 2) put it

"testing reveals the truth".

Even if testing produces many failures, it increases understanding and learning especially in uncertain situations. Small failures can create rapid learning and capture the attention of engineers, so that earlier failures can be mitigated in next iteration. There are cases where physical testing is the only realistic option for evaluation. For instance, testing seals and sealing through CAE analysis is particularly difficult to model because of the non-linear behaviour of sealing distortion. Confidence in a design is reduced when redesigning does not have useful testing results to draw on. Hence, a significant amount of the development effort is spent on physical testing to acquire confidence in product design and decrease uncertainty and risk for the company. In the case study company, engineers are currently more confident in physical testing than CAE analysis. The latter is often referred to as virtual testing. However, for some components, like flywheels, engineers have achieved enough confidence in the accuracy of virtual testing to require less testing physically in early stages of the process. On the other hand engineers in the company also realise the limitations of physical testing. As, physical testing is undertaken with one specific set of conditions, the testing data does not provide information about where the product will fail next or whether the product would have failed if the load was slightly higher than specified. Test results cannot even predict whether another sample part with similar material properties or dimensions, will fail at the same place in the tolerance range.

Further, as the case study company produced engines for different applications and operating conditions the product needs to be tested for each application and set of operating conditions. Multiple iterations in physical testing to cover the whole range of applicability can be prohibitively costly. The company can only afford to test the extreme use cases. Therefore virtual is used to explore the design parameters and the variability of manufacturing parameters, which is not possible in physical tests. In virtual testing, the engineers can systematically vary scenarios, by changing the environmental conditions, feature dimensions and operational values. The cost and time required for these iterations is considerably lower than required for physical testing.

5.1.4 Locations

The testing of a component, sub-system or an engine occurs in many places, including in-house, out-sourced and at a suppliers.

5.1.4.1 In-house

The company has a specialised department called "Global Engine Development" (GED) which has extensive testing capabilities. This testing facility has a large number of engine test cells, and in-house component and materials testing laboratories. The GED facility is to meet requirements for the development, durability validation and evaluation of the products, driven by emissions legislation, customer requirements and business goals. The company considers in-house testing as the cheaper and preferred way with better control over tests. In-house testing facilities enable the design, development and test engineers to work together, resulting in better communication and shared learning.

5.1.4.2 Out-sourcing

The company may need to hire test engineers/experts when they do not have specific expertise. They may also use test facilities elsewhere to conduct a specific test. Out-sourcing happens in several forms:

i) in-house out-sourcing: where experts are hired to assist tests in the company's in-house testing facilities,

ii) partner out-sourcing: objects are sent to another location to use the test facilities of one of the company's partners, and

iii) real out-sourcing: where objects are sent to be tested in other testing facilities. These can happen in both case of CAE analysis and physical testing of product.

Out-sourcing is an expensive option in terms of cost, also might need to commute the objects to test to external test facility. The company makes careful choices of where to out-source, as there are issues in the disclosure of information and the extent to which they will want to share information.

5.1.4.3 Supplier

Suppliers mostly conduct initial component testing. Some suppliers are expected to perform the validation testing for a component or module, for example, an oil pump. The company defines specific software and certain processes of verification and validation to the supplier who validates their product against these criteria. For the example of a pump, which will be used in a specific engine, the company will define the working and boundary conditions and expect the supplier to perform all the durability analysis, reliability assessment based on the boundary conditions. The supplier's product validation testing reduces the level of component level testing in the company dramatically. Access to the supplier testing results and data brings better understanding of the component or module properties and capabilities.

5.1.4.4 Machine testing

Traditionally the company has sold engines, which were stand-alone products. However now, the design of engines, and the design of the vehicles in which they are used, has become more integrated. This is partly because the engine is produced as one output of a global retail group and partly because of industry trends. This increases overall performance while vehicles need to accommodate the heavy after-treatment equipment to complying with emissions legislation. This has led to an increase in machine level testing in addition to the system (i.e. engine), sub-system and component testing. During engine development, this company tests prototype engines by installing these in machines. The company has in-house facilities and external, off-site, testing facilities to perform machine tests, including field tests.

Field tests are especially essential when the company's in-house testing facility can't create the operating conditions, for example, altitude. Engines can also be found running in mines hundred metres below sea level as well as on mountains thousands metres above sea level. The effect of altitude is measured by comparing validation results from the company's different test sites, situated at different altitudes as well as by running field tests at different altitudes.

5.2 A framework for test planning

The previous section describes four entities - object, property, mode and location, which characterise testing activities. These entities can be used to formalise the process for test planning. Figure 5.2 shows these entities in four axis directions/dimensions. A framework for understanding the key steps of testing planning is presented, using the entities that are considered in each step of the plan.

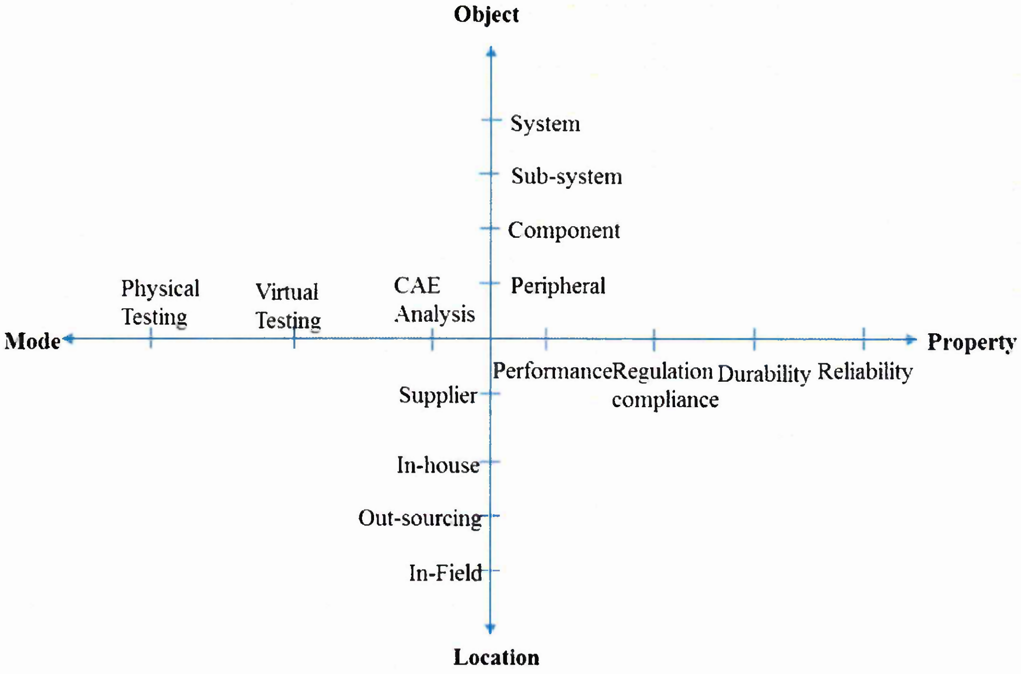


Figure 5.2 Four dimensions that characterise testing

In section 4.2 the test planning process of the case study company was described. This research has identified that there are four main steps in test planning:

i) Risk assessment- identifying design risks through FMEA analysis

ii) Activity planning- i.e. identifying risk mitigation activities and associated modes of testing,

iii) Allocating and Scheduling- i.e. planning those validation activities in stages of PD and in appropriate testing places, and

iv) Validation program planning - i.e. validation program cost and time vs project cost and time.

5.2.1 Step 1: Risk assessment

Planning of testing starts with risk assessment of a product (as a system), its subsystems, components and peripherals connected with it. These are the 'objects' described above. The risk is assessed through FMEA analysis (FMEA is described in Section 4.3.2). A FMEA is attached to an object to assess the risk that given its current capability, it can deliver its required functions in terms of its 'properties'. Critical properties are related to performance, compliance with regulation, durability and reliability. At this stage of the planning, the intention of the FMEA is to identify all the possible ways that an 'object' might fail to meet or deliver it's desired 'properties' and the associated risks of failure. Therefore, the focus of the risk assessment step of the plan is the dimensions of 'object' and its 'property'. In Figure 5.3 , an example is shown, where a component is assessed against all of its properties.

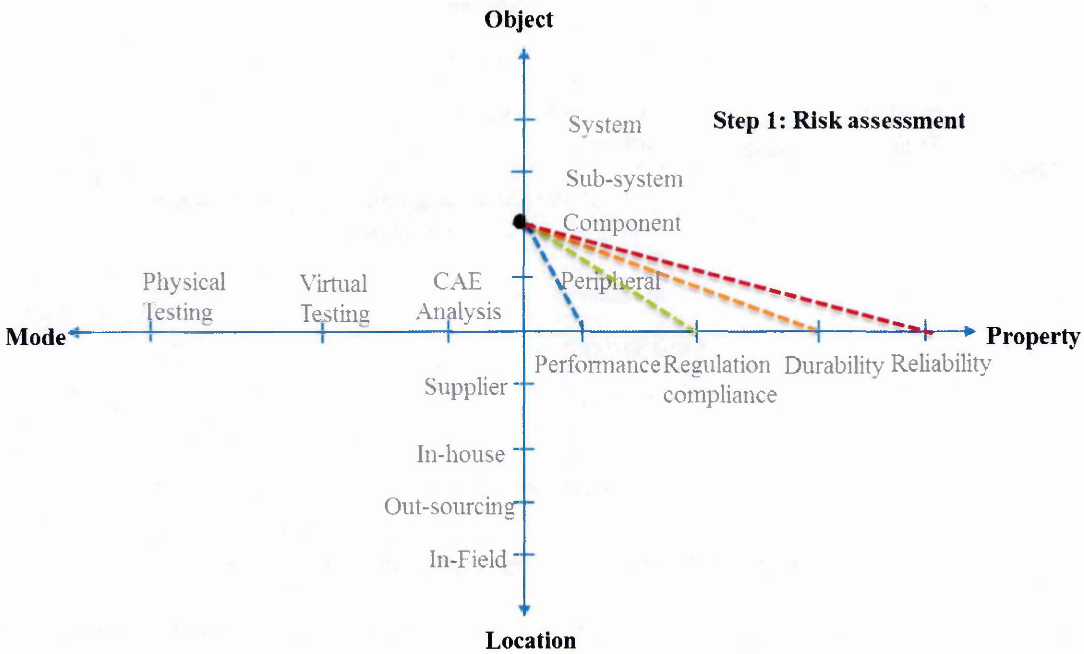


Figure 5.3 Risk assessment of a component by considering all properties

5.2.2 *Step 2: Validation activity planning*

In the next step of test planning, risk mitigation activities are identified. These activities are identified based on the object and the property that is at risk. A testing activity can be a 'mode' of CAE analysis, virtual testing or physical testing. Also an activity, for example, a performance test to measure the fuel consumption, can be performed at least three times at different phases of the product development process; at the System/concept Demonstration (SD), Design Verification (DV) and Product Validation (PV) phases. As the objectives of testing in these phases are significantly different, the focuses on the testing activities in these phases are also different. A combination of modes of activities is used for each phase. These activities are also driven by external factors such as regulations and customer requirements.

Figure 5.4 shows a schematic of planning validation activities on a component. In this case, the performance of a component is planned to be tested in different modes. Similarly, for this component, appropriate activities will be identified for validating other properties, like regulation compliance, durability and reliability. At this step, a complete list of activities is identified for each of the component, sub-system and system level objects. Finally a master FMEA is created, in which all the risk mitigation activities are listed. These lists also contain the duration of each activity and phases of product development where the activity is to be performed. At this level, detailed planning on how testing activities can be aggregated, or disaggregated or combined are considered, in a preliminary form. Several individual physical tests are often aggregated to perform in a single run, as this can save time and setup costs. For example, small parts and components of the engine also go through a valuable thermal cycle test during a durability endurance testing of a cylinder head.

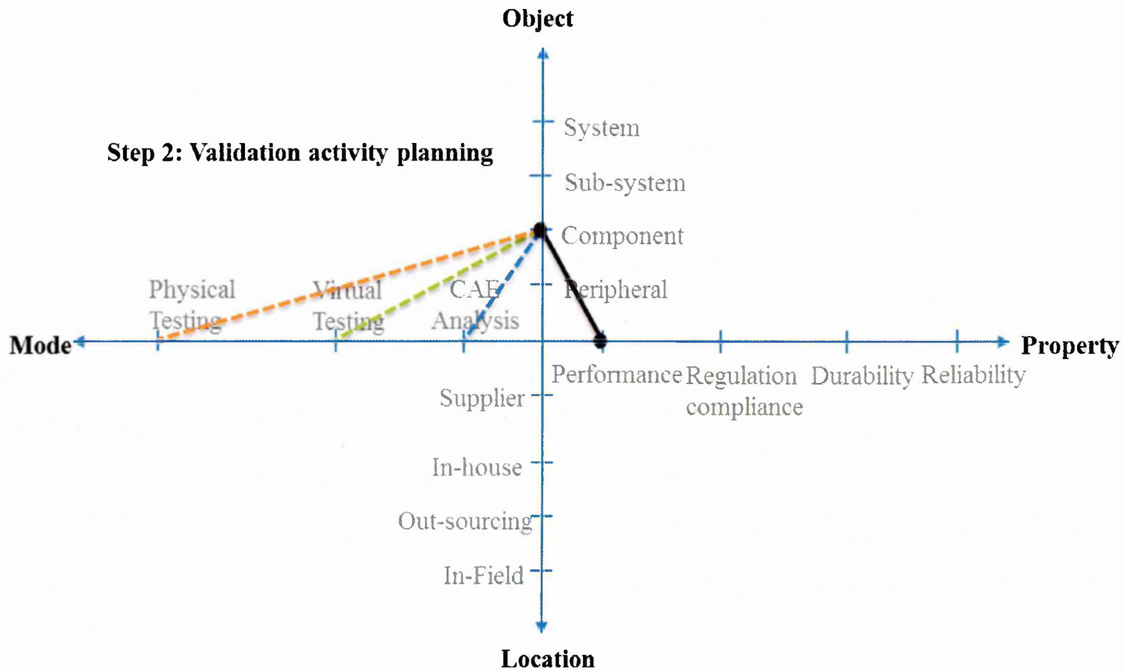


Figure 5.4 Activity planning for performance validation of a component

5.2.3 Step 3: Allocating and Scheduling

Once, the activities are identified and product development phases of these activities are known, the process of scheduling these activities can start. In step 2, all the conceivable validation activities are considered. But, the company cannot viably perform all the activities because of the time and costs that are required. In this step 3, the planning is around choosing the alternative activities and arranging these activities across the PD stages and allocating the testing facilities-i.e. the set-up of these activities is assigned. There are two types of timing associated in this step of planning; 1) when an activity occurs in the process, and 2) the duration of a testing activity. The duration of testing activities is important these must be finished before the next review in the stage gate process.

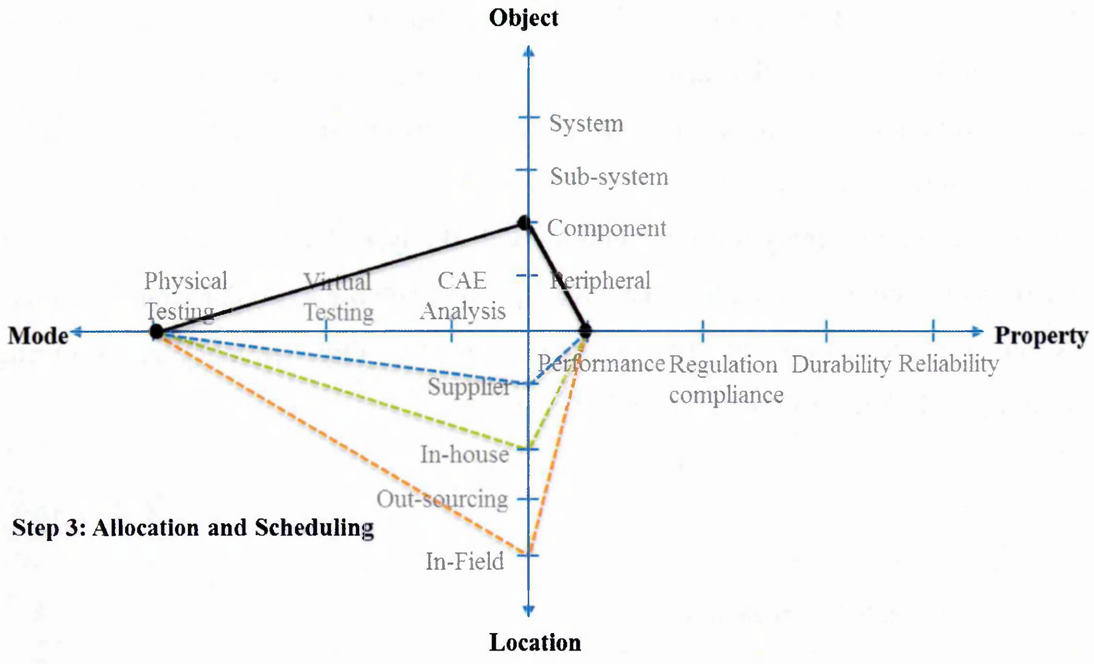


Figure 5.5 An example of allocating and scheduling a physical testing for performance validation of a component

As the in-house testing facilities are shared between parallel projects, the process of allocating and scheduling is largely affected by the 'test capacity plan'. Allocating a test in this step essentially has to be coordinated with the plan of available testing capacity. Alternative locations are also considered in Figure 5.5. For instance, when a supplier physically tests durability of a piston before delivery, the company might not plan to test a piston individually, but will check the piston's durability by incorporating it in an in-house engine test. Further, when an engine is tested in field, this piston will be subjected to a durability test.

The decision on how to configure tests, especially whether to use sequential or parallel testing occurs in this step. When there is no information flow between two tests, then the decisions are based on the durations of individual tests and associated costs. Parallel testing of two lengthy tests can reduce overall testing time but it also depends on the availability of testing facilities. However, some tests must be scheduled sequentially because one test provides input to the later tests. The total cost and time required for the validation programme is calculated in step 3.

5.2.4 Step 4 Planning of validation

The cost and time requested for a validation programme must be within the total cost and time allowed in product development. In this step, the decision is taken as to whether the validation programme will achieve the required reliability at each stage and that it will be within acceptable time and cost. The example in Figure 5.6 shows validating performance of a component through physical testing.

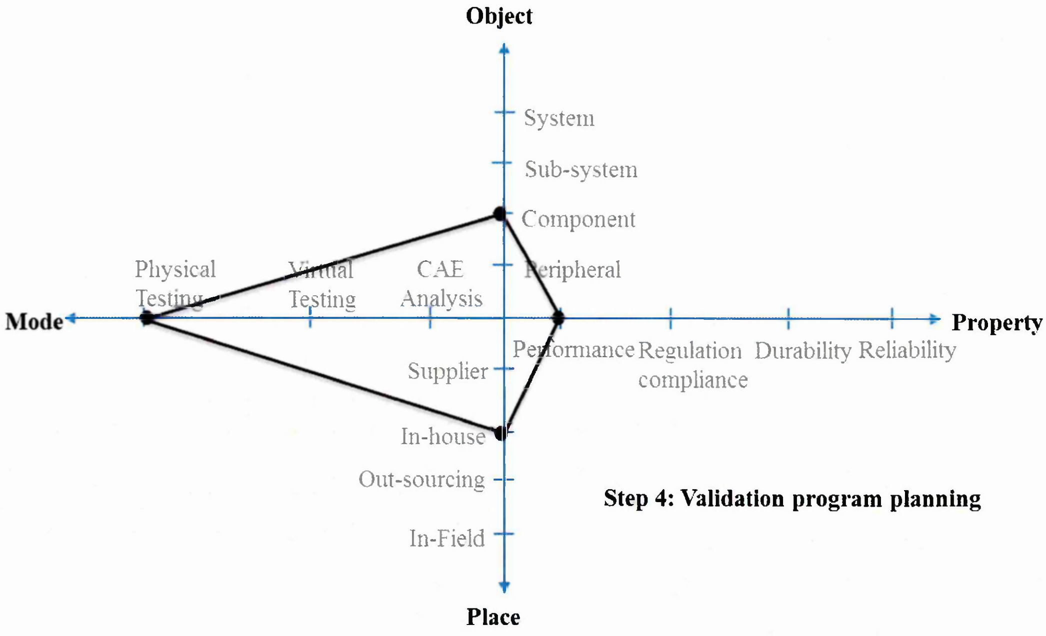


Figure 5.6 An example of validating performance of a component through physical testing

Planning validation is a balance between "reliability can be achieved" and "the cost and time of that programme". For example, a highly reliable validation programme might incur significant costs and alternative validation activities may need to be chosen if the cost and time are beyond the budget of the programme. Any changes in validation activities require reiteration of the whole planning process, i.e. the steps of activity planning, allocating and scheduling and validation program planning.

The section above describes the whole process of testing planning and those entities that are considered in each planning step. Figure 5.7 maps these steps into the company's process of testing and activity planning, (previously shown in Figure 4.7). Along with the four dimensional framework described above and by mapping these steps into the company's process, it can be clearly identified the reasons of each steps and the entities to consider in each step of testing planning. As mentioned previously, the process of testing planning starts early in the product development process, almost at the same time as project planning. These plans constantly change as the product development progresses.

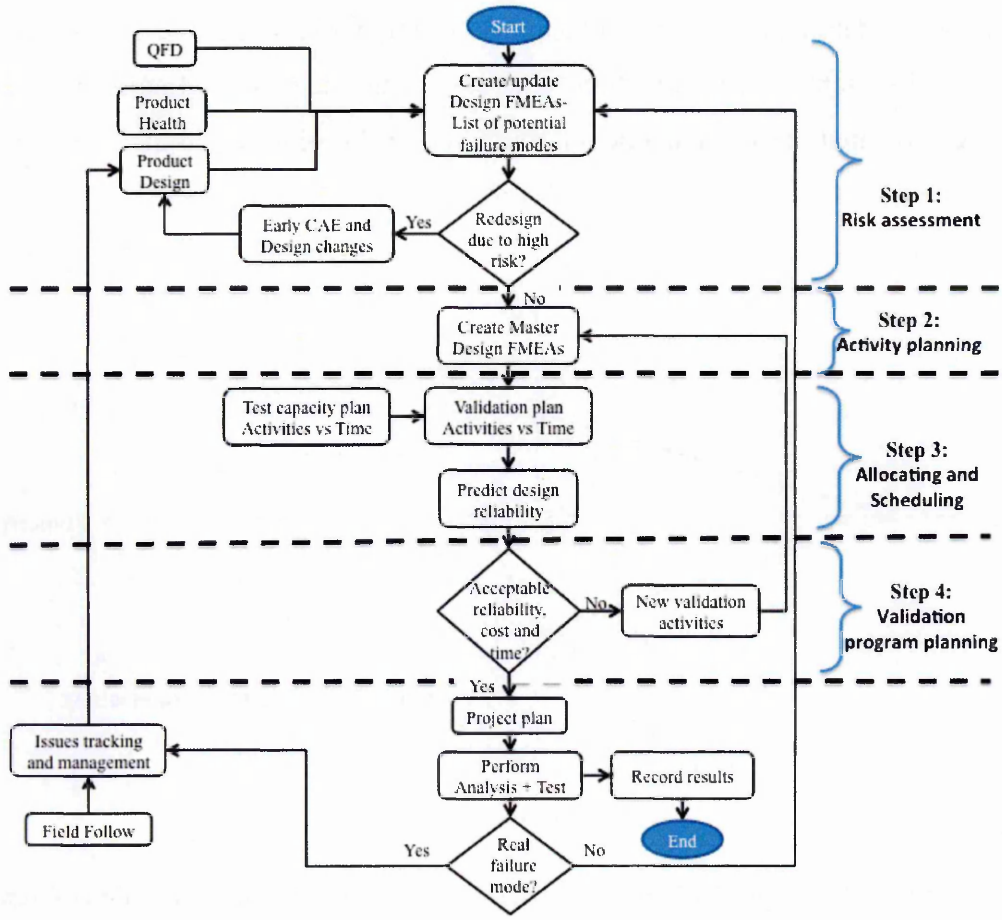


Figure 5.7 Steps of testing and validation activity planning

5.3 Effects of testing on the product development process

In the previous section a framework for test planning was mapped onto the overall flow of activities in product development (Figure 5.7). In this section the testing and design activities are mapped onto the product development gateway stages (see Figure 4.5) and analysed to illustrate the effects of testing on the product development process.

Figure 5.8 presents the structure of product development activities that was established through intense analysis of the process structure of the case study company 1. For simplicity, Figure 5.8 presents the key activities as time limited boxes but in reality, a core team keeps working on Design and CAE throughout the entire period, and testing goes on almost continuously, in parallel to these activities. There is a fluctuation in effort through the period, as resources are shared across the different engine projects, which can be at different stages of development. For example, the design team will be heavily involved during GW2 of a project and will be maintaining any design changes in GW3 and GW4, while working on GW2 for another project.

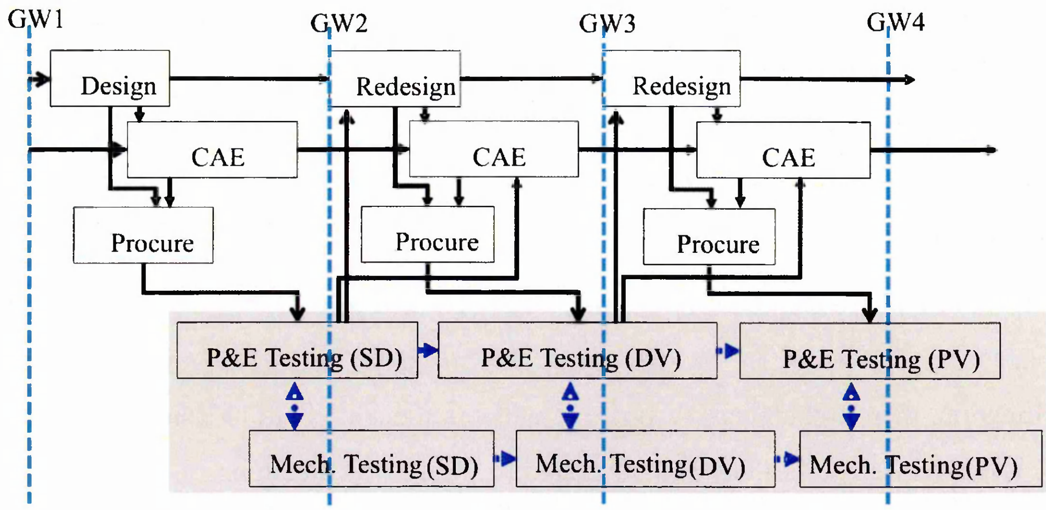


Figure 5.8 A schematic of the product development activities from GW2 to GW4 (SD = System Demonstration, DV = Design Verification, PV = Product Validation, P&E = Performance and Emission)

Another representation indicating how iteration between design, virtual and physical testing takes place at each stage of product development is shown in Figure 5.9.

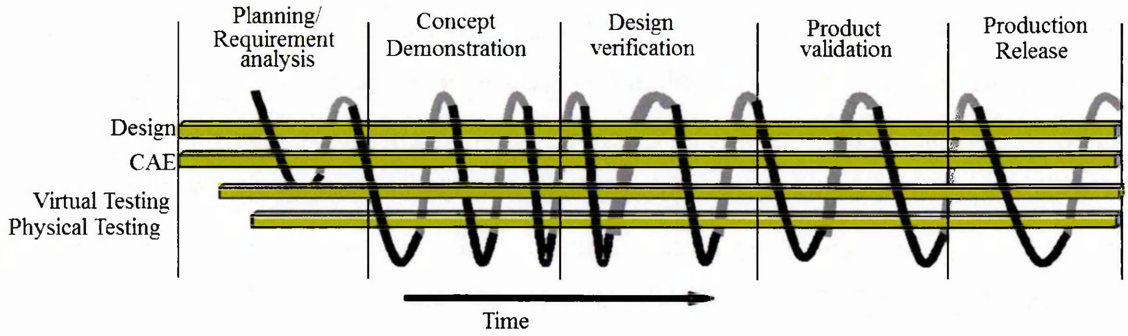


Figure 5.9 Product development integrating iterations between Design, CAE, virtual and physical testing at each stage

This structure of activities is significantly different to that found in the literature, because it closely interlinks testing activities with design. There are two main effects of this interlinking. First, design and testing are essentially iterative and second physical testing can cause process delays.

5.3.1 Iterative nature of testing and design

Section 4.1.3 illustrated how engines are designed and tested in sequence for system demonstration (SD), then design verification (DV) and product validation (PV). However, in reality, several versions (usually three variations) of the same engine are tested simultaneously in parallel test-beds, where each bed replicates a particular set of specifications, typically the worst-case scenarios. Stages frequently overlap, so some components are tested for SD whereas others are tested for DV. As noted above the process described in Figure 5.8 is essentially iterative and one (re)test leads to the iteration of (re)design and vice versa.

Testing in one phase can identify design issues and lead to re-design in the next phase. For instance, if testing in the SD phase identifies a failure or mismatches with specification, then in the next DV phase, engineers focus on both redesign to overcome those issues as well as further detailed design for design verification.

Emerging design changes can lead to re-testing and changes in future testing plans. As every part of an engine has complex connections with other parts, design changes in one part can cause changes to its connected parts which might effects at sub-system or system level. Therefore, a design changes can propagate, which nullifies some of the existing testing, introduces more testing and questions whether previous testing was adequate or was performed in a right way. For example, if a component fails to perform specification in the DV phase, engineers will improve the design of the component while analysing how those changes might affect other components or the performance of the whole engine system. The validation manager will require test to be planned both for that particular component and for affected components. Engineer might not necessarily perform the same testing activities as in previous stage but incorporate new testing parameters. Re-testing might happen in different mode, for instance, virtual testing might be enough to verify a design change and physical testing might not be necessary. However, major changes in design require new system level physical testing and this can delay product development.

5.3.2 Physical testing causes process delay

Analysis of the case has suggested that there are three main effects of this delay: missed gateway schedules, mismatches in test-bed schedules, and accumulated delays.

5.3.2.1 Missed gateway schedule

The company has a strong emphasis on maintaining each gateway using gateway-reviews for assessment and monitoring. At each stage, testing activities are scheduled in such a way that the gateway timeline can be maintained. Some activities might have flexibility, for example, engineers might decide CAE analysis could be finished before the planned timescale as they have already achieved the desired output. On the other hand most of physical testing are restricted to planned time-frames. If a physical test is designed to run for 1000 hours, that test must run for that specific time, unless a failure occurs before that. Even if a failure occurs, engineers will likely replace the failed component and continue to run that test to learn about other components' behaviour and durability and complete the test. Therefore, in the case of a physical test starting later than planned there is little chance that this test can be shortened. The consequence is that causing the activities at this stage will fail to meet the gateway schedule.

5.3.2.2 Mismatch in test-bed schedule

As the company shares testing facilities across several projects, the validation manager plans the tests and allocates the test-beds very early in the process usually during stage 1 and stage 2. If a test-bed is occupied longer than planned then the next batch of tests is disturbed and test-bed schedules are mismatched. A catastrophic failure at engine level can cause a test to stop and the test-bed can be unoccupied for the length of that test. Depending on the priority placed on finishing a particular project, the testing of a new component might occur urgently, and disregard scheduled tests. Consequently testing might not only affect the next stages of the project, it can also disturb other project's schedules.

5.3.2.3 Aggregated process delay

Delay in testing activities in one phase can affect the related activities in next phases, with consequential delays in next phase testing. As a result, delays aggregate and cause overall process delay. This can lead, eventually, to late time-to-market.

5.3.3 Issues with physical testing

Physical tests are usually time consuming and involve many facilities and personnel. For instance, the 'gross thermal cycle' is a 1000 hour endurance test, which ascertains if the cylinder head will endure repeated gross thermal cycling without cracking. Usually this test is performed at the DV phase. The test involves a fitter, a tester for full time and an engineer for about 20% of his time.

In analysing the company's design and testing processes, two issues were highlighted. First, long lead time for procurement and second the long duration of physical tests. As physical testing needs physical objects, time is required for procuring the parts, building the engine, testing and analysing the test results. For a durability test, 'gross thermal cycling' for instance, the company needs three months to procure the components, two weeks to prototype engine building, actual testing requires two months and another two weeks of post-processing. In total six months are required to complete this test and develop a finished product specification.

5.3.3.1 Long lead-time for procurement

The lead time referred to here is for procurement of items or prototype systems for testing. It is the time between the placement of an order and delivery of an item from a supplier. As physical testing needs physical objects, the company needs to procure prototypes or production intended objects from the supplier. The company allows at least three months for core components, like – the cylinder head and cylinder block. Any delay in placing the order increases the lead time. Further, any change in the specifications can significantly increase the lead time.

It was observed that, in many cases, when the company needs to start the test to meet the schedule of the next GW stage, an important hardware component might not be available from the supplier. The company cannot afford to delay, and instead tests using alternative off-the-shelf components or makes the prototype components in different ways, e.g. a component might be machined for prototype that later will be cast. The validation manager identifies suitable alternatives and calculates trade-offs. For example, to test the cylinder head, which, will not be delivered until later, the engineers will either continue physical tests with a prototype cylinder head or simulate the engine computationally and identify the associated risks. However, these alternative tests may not provide planned levels of risk reduction. In this scenario, the product cannot be signed off, and physical testing of the cylinder head is still necessary for verification. This situation causes the DV or PV phases to extend over two Gateway (GW) stages instead of one.

5.3.3.2 Lengthy physical tests

Usually reliability tests are extremely lengthy, because these tests are designed to predict the life-time behaviour of the products. Lengthy physical tests like 'gross thermal cycling', as mentioned before, is a validation test for determining the thermal fatigue resistance of core engine components by putting an engine on a test bed for 2 months in a stressed condition. Also this test needs to run for at least three times for three different extreme use cases. That means three engines are required to run in these three different specifications of this test. These tests can be run in parallel if the facility, i.e. the test beds and the personnel are available. These tests are essential and mandatory but are very expensive. Therefore, there can be a tension between the engineers about these types of testing. Some engineers think these are highly valued tests and they can learn a lot from them. But others think that as these tests are already in a stressed condition, the time for the test could be reduced by increasing the stress conditions, and deliver the same information. The latter is unlikely to be the case because there is a possibility that the engine might fail or break down if the stress is increased. The company generally performs non-destructive tests and follows the overall procedures and guidance of the business which involves lengthy physical tests of use cases.

As lengthy procurement time disturbs the process and some physical testing takes a long time, the DV phase testing may still be on-going while the (re)design for the PV phase is started and while procurement for the subsequent PV testing begins, as seen in Figure 5.8. Without the testing results being available, there will be uncertainties in redesigning and procuring for the next phase. This uncertainty can result in increased number of design and testing iteration in subsequent phases.

5.3.4 Company strategies

The company is currently counteracting these problems with different strategies. Procurement issues are tackled through good and frequent communications with suppliers; physical testing issues are moderated by supportive CAE analysis and overall process delays are minimised by re-arranging internal activities.

5.3.4.1 Strategies for procurement issues

Initially the company tries to provide accurate product specifications to the supplier and freezes design continuously by maintaining good communication with supplier.

Better specification to the supplier

To minimize long lead-time procurement, a clear and accurate specification of the product is required. The company uses CAE analysis and makes virtual prototypes with many iterations to enable the first physical prototype to be built closer to target. Engineer 1 commented,

> "computer simulation is becoming increasingly important to the companies to minimize the effort and expense involved in product development".

CAE analyses enable the company to carry out optimization earlier in the product development cycle (front loaded), as well as improving product specification to the supplier. Clear, precise and accurate specification of product and testing requirements help supplier product development and validation process and can reduce the procurement time (as mentioned by Engineer 2).

Continuous design freezing

Ideally, design engineers will work on detail drawings, specifications, and instructions and product engineers will perform all engineering work on the final design. When these tasks are completed and signed off, the final design is frozen and released for procurement. Design changes are restricted after the final design freeze. But the company employs continuous design freezes. Component specifications are initially released to the supplier based on the CAE analysis as a basis for actions and occur as early as possible. These specifications are continuously improved thorough design iterations and detailed CAE analysis. These design specifications are also developed and changed based on collaborative information exchange between the company and its suppliers. The company's continuous design freeze process allows the engineers to react to design changes ahead of the final design freeze. This can reduce the time and resources required to deal with quality problems.

Good communication

Effective collaboration allows the company to use the knowledge and expertise of suppliers to complement internal company knowledge and capabilities. Suppliers became active in exchanging problem and solution-specific information with the company. Good communication with supplier enables the company to get access to information about any design issues promptly; as a result the company can perform any design changes with suppliers validated results. Further, good communication can provide possibilities for negotiation about changes and modifications. Good communication is considered a key factor to the company to reduce the procurement time of prototype components.

5.3.4.2 Strategies for physical testing issues

The company tries to deduce the duration of a lengthy physical tests by accurately designing the test setups before the test starts and by reducing physical testing effort during the test through substituting physical testing activities.

Accurate test setup

Before starting a physical test, CAE modelling and simulation drives test requirements and engine settings for the test. CAE engineers together with test engineers spend lot of time in analysing test settings to accurately provide specific information for required variables, for example, sensor locations or sensor and system calibrations. These analyses are also performed to recommend the expected values for test scenarios. For instance, load cycle and loading locations can be recommended to maximise information derived from tests. These activities can considerably reduce the testing effort and the test engineer can provide feedback to design more reliably.

Activity substitution (replacing physical testing by virtual testing)

The company can only afford the resources to perform physical tests for a selected set of specifications, which are usually worse case scenarios, although company's engines are used in thousands of different applications. However, design variations for different applications are verified and validated through virtual testing. For example, structural integrity analysis, including fatigue prediction under various loading conditions, e.g. static and dynamic stiffness analysis and transient and vibration analysis; are evaluated virtually by simulation for all intended application. These virtual tests have substituted and significantly reduced the number of prototypes required to be built and physically tested. Even during a physical test, test engineers are provided with predicted value and behaviour of the product under test, so less time is spent on physical testing.

5.3.4.3 Strategies for accelerating product development

Frequently, lengthy physical testing and long procurement times can be the cause of a stage delay and the result is process delay. When the company fails to maintain the planned schedule in stages, engineers decide to accelerate the process by concurrent execution of activities and increasing the intensity of the lengthy activities. Effectively they use more resources to finish on time by overlapping activities (concurrency) or intensifying Activities (crashing).

Activity overlapping

The process of starting an activity before finishing the downstream activities is called overlapping (Krishnan & Bhattacharya 2002, Gerk & Qassim 2008, Roemer & Ahmadi 2004); thus overlapping of two previously sequential activities can save time. Ideally testing of one phase should be finished before design of the next phase can be started. In Figure 5.8, it is prominent that design activities are starting before finishing the testing of previous stage and this is happening in every stage, hence two sequential stages are overlapped to maintain the gateway schedules. Companies often also overlap lengthy physical testing activities to minimize the total duration of testing. So there are two types of overlapping: activity overlapping and stage overlapping.

Activity crashing

Activity crashing is a process in which an activity is shortened (crashed) by increasing the intensity of work of that activity (Gerk & Qassim 2008, Roemer & Ahmadi 2004); consequently subsequent activities can start earlier. Engineers consider that the company's process is already designed with full intensity of work and there is very little scope of activity crashing. However, emergencies, where there is a chance of project delay, the company might decide to increase the intensity of a testing activity by incorporating more staff, working on weekends, or hiring external experts. They might also send a prototype component to external testing facilities to finish a physical testing. Often other tasks such as CAE simulation and modelling can benefit from activity crashing. Simultaneous execution of simulations in parallel might finish the job quicker and subsequent activities can start earlier.

Activity crashing has little value in some cases. For instance, some tests are designed to run for a specific length of time, for instance, if a durability test is designed for running an engine for 1000 hours in a test-bed, the engineers have to run that test for that long. Increasing the effort cannot reduce duration of these kinds of tests. Also, activity crashing is an expensive approach as the efforts are increased, for instance, more personnel are included and more tangible and intangible resources are incorporated to finish the activity quicker than initially planned. Sometimes this process can disturb the existing schedules of the staff. Activity crashing might urgently bring an individual into an activity who was scheduled for another critical activity at that time.

The issues of physical testing, especially, the ways that these testing activities are distributed across the phases of product development reveals a complex picture highlighting several significant research challenges. The next section sets out these challenges and the remainder of the thesis is devoted to developing methods and tools to address these challenges.

5.4 Emerging research challenges

The findings from the main and subsidiary case studies have provided a better understanding of the testing process in industry and contributed towards answering initial research questions. This section reviews these questions and then sets out the challenges, which have emerged from the case studies.

5.4.1 Answer to initial questions

The findings from the studies in Company A and Company B contribute to answers to the initial research questions posed in Chapter 1. The first question is "*how testing is integrated into the product development process?*". The analysis of product development in both companies supports that testing process starts very early in the process, even before the detailed design of the product starts. This early stage testing can use modelling and simulation of concepts through CAE analysis, to explore the design opportunities and alternatives. Subsequently, testing is performed on a preliminary design, which is a combination of new and old components to check a product's capability to meet the required performance. Design verification testing is executed next to develop the performance and hardware validation prior to commitment to expensive production tooling. Finally, product validation testing checks the effect of production variability on performance as well as any remaining hardware variation. Testing and design are closely intertwined and interrelated from the beginning of the process and in each phase of the product development.

The second question is "*what are the roles of testing?*". The different purposes of testing have been discussed. Traditionally, testing has been considered as tasks to be performed to validate the product and for acceptance testing. However, the companies perform tests for many more purposes, such as for learning, demonstrating, concept selecting, design verifying and confirming. Perhaps, most importantly for the product development process, testing gives the confidence to the engineering activities. The role of testing also varies in different phases of product development. In the early phases, testing is for learning, demonstrating, selecting, and in the later stages testing is more focused on verifying, validation and confirming the design.

Testing can be performed in different modes- through CAE simulation and modelling and physical hardware testing. Different modes can give different levels of confidence and vary between companies. In both case study companies engineers feel more confident in physical testing results than CAE modelling and simulation. However, CAE plays a vital role in shaping the boundary conditions by specifying the expected outcomes of a physical test. On the other hand, physical testing also validates the CAE models and simulated results.

Testing is one of the major drivers of the design process. At each stage of product development in both companies, testing provides the confidence and confirmation to the engineers required for sign-off to proceed to the next phase.

The third question is "*how these testing tasks are scheduled across the stages of the product development process?*". Risk is the factor that drives the testing planning in these companies. At the start of the project, risk is assessed based on current understanding of the technology, product and processes. New or unproven technologies are considered as high risk because of uncertainties associated with them. Also any design changes raise uncertainties to the system. Both companies use a tool called FMEA to measure the risk in the project and assign tasks to mitigate risks. These companies realise that all the risks cannot be removed within the cost and time limits of a project. The tendency is to remove the high risks and mitigate the others. Products can be high risk when they do not comply with legislations and cannot be marketed. Therefore, the companies' test planning is significantly focused on verifying and validating the product to comply with legislation.

Testing planning is complex in nature since many aspects of planning are combined. Planning of a single test includes four factors; the properties, which are tested, the object under testing, the mode of testing and the location of the test (discussed in section 5.1). The property to be tested usually indicates the phase of product development at which the test is required but this can vary. In both companies, test planning starts at the beginning of the project, because physical testing needs to secure testing facilities and resources. These resources include, test beds, physical prototypes, test engineers and other facilities. CAE modelling and simulation also require planning for new tools and applications, training of engineers and managers in using these tools, and their allocation to project tasks. As testing resources are shared between multiple projects, which run in parallel, there is a tendency to secure these resources as early as possible. However, this leaves many uncertainties in place.

In company A, the testing planning is according to the standard of the total business. Company needs to follow the guidelines that have been practised for many years. Engineers have different views in these guidelines and procedure of tests. Some senior personnel think these guidelines are out-dated and require to be improved. But the business is reluctant to change because these guidelines have proven records, and changes can raise uncertainty into the process; also not everyone involved in these testing process shares the same view of changing.

There are three major challenges which are outlined in this section and then considered in detail in the following three chapters:

i) Prioritising testing activities

ii) Scheduling testing activities

iii) Overlapping testing and design activities

5.4.2 Prioritising testing activities

In both companies, current testing planning is based on FMEA analysis. Testing activities are ranked based on high risks identified in FMEA. However, as the complexity of the product increases, potential failure modes also increase. With the limited resources available, it is not viable to test all the failure modes, it is therefore important to first prioritise the failure modes that need attention for testing.

When in FMEA, all the risks are listed; the related risk mitigation activities are attached with each of the risks. The key role of the validation team is to manage these activities to make sure that they accept the right risk deduction verification and validation activities and reject the unnecessary ones. The validation team will also see if requested activities can be carried out within budget, are within an allowed time frame. These selections of verification and validation activities are based on subjective judgements of engineers and mostly driven by the critical areas highlighted in the FMEAs. However, there is no tool or methodological approach for selecting Design Verification (DV) and Product Validation (PV) activities.

This research has also identified that the construction of FMEAs relies on understanding of the product and current issues with the product as well as available engineering expertise and intuitions. FMEA owners can have little knowledge of initial customer requirements. Therefore, customers' voices are not effectively included and considered during the construction of FMEA. The understanding of the customers' voices and their effect on the design soon starts to blur in the process of FMEA creation. In the detailed FMEAs, at system and component levels, there is no indication of customers' voice and their importance is rarely considered, mainly because there are two different teams dealing with QFDs and FMEAs.

QFD teams translate the critical customer requirements (CCRs) into product objectives and technical characteristics and hand them over to FMEA teams. The QFD teams understand the importance of customers and the FMEA teams realise the technical aspects of the product. Often these teams do not share common understanding and knowledge, partly because there is little overlap of membership between these teams. The validation manager, who mainly deals with FMEAs, understands which technical requirements are being measured in a particular test. However, often there is a gap in understanding of the critical customer characteristics they are being satisfied through that validation test.

The voices of customers are important drivers for testing (as mentioned by the senior engineer –Engineer 1) but these are not directly considered during the construction of the FMEA. Although the company uses quality function deployment (QFD), in capturing the customer requirements, there is no systematic method of integrating QFD and FMEA for ranking testing activities.

Ranking and prioritising testing activities are important because there can be situations when the number of tests are required to validate the product is significant and the cost of performing these tests well exceeds the limit of the budget. By prioritising the testing activities, product engineers and the validation team can focus on testing activities, which are mandatory and essential for product validation and legislation purposes as well as for customer satisfaction. Less important testing activities can be dealt with alternative testing methods, which might still give them sufficient levels of confidence in the validation.

5.4.3 Organising testing activities

As pointed in the literature and in the company, tests are linked in such a way that one test may require inputs from other tests or can prove other tests inadequate or invalid. Thus, a systematic way of representing dependencies between tests is vital for effective test planning. With testing being so expensive in time and resources, it is critical that design process planning deploys test activities effectively throughout the product development process, taking into account dependencies. However, a typical validation plan rarely includes information about any dependency and interrelationships between different tests. In the current process at the case study company, the analysis of FMEA generates a plan called the design verification plan and report (DVP&R), which includes a list of validation activities, starting and finishing dates, phases of the PD process to be performed, which item of product, and what parameters to measure. These lists are maintained and managed in a spread-sheet. It can be difficult to visualise the dependencies in this form. There is no supporting tool for identifying or connecting or visualising these different types of testing activities – physical and virtual – scheduled in the product development process.

5.4.4 Overlapping with uncertainties

In Section 2.4.2, the issues with overlapping tasks and stages of product development were discussed briefly. Engineering companies overlap testing and design as essential practice. The case study company overlaps the tasks and phases of product development that results in many uncertainties in the process and can cause overall process delay. The company has no choice but to overlap design tasks with testing, as a design proposal is needed to commence another, often lengthy, procurement process. A way of accelerating the testing process was critical. There is little academic research that has investigated the parameters that affect the overlapping of testing and design tasks, especially in the context of a complex product development process.

5.4.5 Integrating virtual and physical testing

The case study company currently uses CAE analysis and virtual testing as a strategy to minimise the issues with lengthy physical testing.

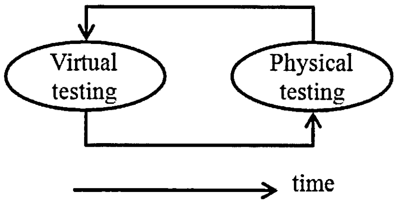


Figure 5.10 Current sequential iterative views of VT and PT

In this current sequential approach (as in Figure 5.10), physical testing is supporting to create virtual tests and then virtual tests are useful for finding boundary conditions and to set-up physical test conditions, input parameters and sensors locations. But this approach does not assist to reduce the physical testing time or can't replace the need of physical testing. Virtual models are usually calibrated with previous test data and are used to predict future. Often this is a reason why virtual testing might not replicate a product's actual behaviour that occurs in a physical test. There is a need for improved concurrent integrated approach of physical and virtual testing.

Chapter 6 A tool for prioritising testing activities

The empirical study described in Chapter 4 identified a strong need for prioritising the testing activities. Also, Chapter 5 (especially Section 5.4.2) highlighted the issues with current approaches in the company. This chapter describes a method for prioritising testing activities that was developed to address and resolve the issues that are currently present in practice.

Section 6.1 outlines a brief introduction to the key factors that were used to develop the method. Section 6.2 describes the method and a tool for representation. An example is used in Section 6.3 to illustrate the method. The evaluation of the method is presented in Section 6.4 and a summary of the key ideas are discussed Section 6.5.

6.1 Prioritisation factors

A product, system, module or component might fail to meet its functional requirements or technical objectives (TOs) and there can be several modes of failure. As the complexity of products increases, the number of potential failure modes increase. With the limited resources available, it is important to first prioritise the testing activities from the both a technical engineering and a customer's point of view. Prioritisation helps to identify the most valuable tests from a set by distinguishing the critical ones from the trivial ones.

In traditional FMEA, a risk priority number (RPN) is computed for each failure mode. The failure modes are subsequently prioritised by RPNs. Efforts to prevent and detect those failures are focused on the failure modes with the greatest RPN values. Testing comprises activities for detecting the failure modes of a functional requirement or technical objectives. Currently, the prioritising of testing activities is based on the high risk areas or high RPNs, identified in FMEAs. Engineers use the FMEA tool for guidance, to predict "*where the program is*" in terms of risk.

The company also relies heavily on engineering judgement and experience. They decide the serious failure modes and perform the tests whether the FMEA indicates these high risks or not. Engineer 3 commented,

> "the expertise suggests focusing on the most recent and major issues, but the FMEA is useful exercise and should be highlighting the unconcerned issues, which can be ignored otherwise."

Besides detecting failure modes, testing activities also ensure that the design is verified and that the product is validated according to specific customer requirements and compliance with the regulations currently in place. Although, based on the importance of customer requirements, Technical Objectives are prioritised in QFD analysis rather than being considered in FMEAs. As Engineer 3 mentioned;

> "Failure modes, we don't weight by customer requirements, we weight them by the risk of occurring"

However, engineers also realise that important and critical Technical Objectives should receive more attention than the less critical ones. Therefore, the importance of testing activities should also depend on the importance of individual technical objectives.

Nevertheless, prioritising a testing activity should measure how important a test is compared to other tests in the group. Typically, a system level test verifies and validates multiple technical requirements or objectives. For instance, a 'heavy duty endurance cycle' test will evaluate maximum cylinder pressure, in-cylinder temperature, fuel consumptions of an engine and so on. Also, a technical objective is evaluated through multiple testing activities, for instance, maximum cylinder pressure will be measured and evaluated during 'thermal cycling' and 'fatigue analysis' tests. Therefore, although one test can be significantly important to ensure a specific requirement, another test should get similar attention if it ensures that multiple requirements have been met. Hence, the relative importance of a test is an important measure during prioritising of the testing activities.

Therefore, this study proposes that the prioritisation of the testing activities should take into consideration the following facts:

(i) The importance of technical objectives based on critical customer requirements.

(ii) The risk of failure to meet those technical objectives

(iii) The relative importance of a test.

6.2 A tool for prioritising Testing Activities

The company uses House of Quality analysis (HoQ) for mapping Critical Customer Requirements (CCRs) to Technical Objectives (TOs). This House of Quality mapping is used to prioritise these technical objectives. FMEAs are used for quantifying the effects of failing to meet those technical objectives. Using these values from the House of Quality analysis and FMEA (as shown in Figure 6.1), this research proposes a method of calculating the importance of testing activities, so that testing activities can be prioritised.

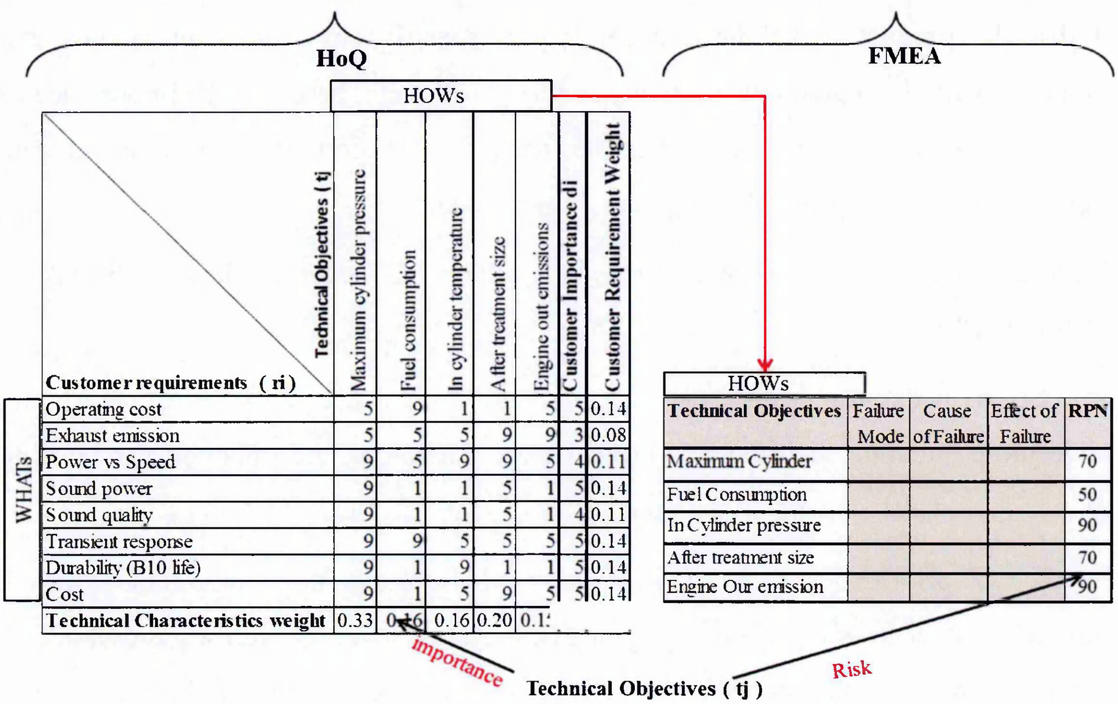


Figure 6.1 Shows the data used from HoQ to FMEA

In the following sections, first, a tool is presented that accumulates information from various other tools and represents this information. Then, mathematical equations are presented for analysing and finally the method of populating the tool and analysis is presented.

6.2.1 A representation of the tool

While the proposed method uses the information from the HoQs and FMEAs for prioritising testing activities, this study also introduces a matrix for mapping the relationship between Technical Objectives to Testing Activities. In this section, first, this matrix is introduced and then the arrangement of the data from HoQ, FMEA and the matrix is presented.

6.2.1.1 The Technical Objectives and Testing Activity Mapping

This is a rectangular matrix is shown in, named Testing Objectives and Testing Activity (TOTA) matrix.

The TOTA Matrix

Technical objectives are listed as the rows in the matrix (Figure 6.2) and Testing Activities as the columns (or can be vice-versa). The cell values represent the relationships between technical objectives and testing activities.

Technical Objectives \ Testing Activities	Heavy duty endurance cycle	Thermal cycle	Emission measurement	Transient machine cycle	Virtual thermal mechanical analysis	Virtual fatigue analysis	Noise and vibration measurement
Maximum cylinder pressure	5	5		5	5	5	3
Fuel consumption	5		5	3			
In cylinder temperature	5	5		3	5	3	
After treatment size	3		5	5			1
Engine out emissions	1		5	3			

1, 3,5 Based on
Correlation between
Tests and TOs

Figure 6.2 A matrix for mapping the Technical Objectives to Testing Activities

Method of modelling the relationship

If a test evaluates a technical objective, then there will be relation between that test and technical objective. These relations can be represented in several forms. In the first instance, it is easier to capture the binary relation, i.e. if there is an existing or non-existing relation. Existing relation can be noted with an 'X' and non-existing relation with a '0' or left blank. In the next stage, the strength of the relationship is expressed qualitatively, for example, as 'high', 'medium' or 'low'. Eventually the strength of the relationship is assessed numerically, such as 5, 3 or 1 – this value of the relationship is assessed by the engineers, who understand these relations and the strength of these relations. A larger value represents a stronger relationship.

6.2.1.2 Arrangement of HoQ, FMEA and TOTA

The arrangement of data from HoQ and FMEA and the TOTA matrix is shown in Figure

6.3. Three boxes are shown: (1) From HoQ, (2) From FMEA and (3) the technical objectives and testing activities Mapping Matrix. Taken together this analysis forms a Test Prioritising Tool (TPT).

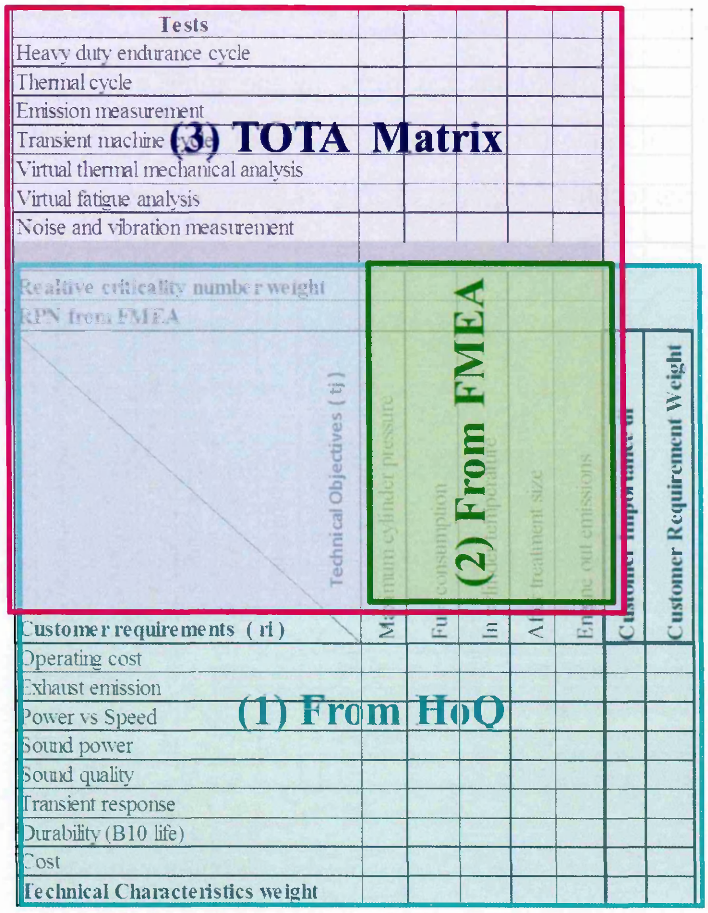


Figure 6.3 Shows the mapping of HoQ, FMEA and TOTA.

(1) *from HoQ:* a part of the House of Quality (HoQ) is considered, without the top roof. In this, critical customer requirements (CCRs) are listed to the left of the rows and technical objectives (TOs) are above in the columns. The intersection cells represent the relationships between CCRs and TOs, The bottom row and the right two columns are used for representing the analysis of this HoQ (will be discussed next). Typically, these TOs from HoQ are transferred to FMEAs and analysed.

(2) *from FMEA:* this captures the results from FMEA analysis, i.e. the associated value of RPNs for these technical objectives are collected and is presented.

(3) *The third matrix is the TOTA matrix,* which is used for mapping the testing objectives to the testing activities. As technical objectives are the shared focus of these three matrices, the TOTA matrix is transposed to fit into this structure.

6.2.2 Mathematical modelling for analysis

This section discusses the assumptions and notations that are used for analysing the data collected in the TPT tool. As this tool integrates the output of HoQ, FMEA and TOTA, the assumptions used in these individual tools are given first and the overall equation is given for analysis.

Assumptions and notations used in HoQ

The analysis in HoQ differs in different companies. Some identify the qualitative relations and others focus on quantitative relations. It is assumed that if a quantitative analysis is practised in a company, the weighted mean of Technical Objectives will be readily available from HoQ and can be used in this tool. The case study company uses a quantitative analysis but it was not possible to present the exact mathematical equations used in the company for confidentiality reasons. Therefore, this research uses the standard mathematical equations for analysing the HoQ. Details can be found in (Hauser & Clausing 1988, Park & Kim 1998). The aim of this analysis is to identify the critical importance of the TOs. The steps of analysing HoQ to find the weighted mean of the TOs are given below:

(1) Critical customer requirements (CCR), are denoted by r_i with index $i = 1$ to I.

(2) The degree of importance d_i, of each CCRs is presented by an integer value from 5 to 1. When the degree of importance of the customer requirement is high, d_i assumed to be 5; and 1 for low.

(3) The weights of the CCR are denoted as $w_r(r_i)$ and defined as:

$$w_r(r_i) = \frac{d_i}{\sum_{i=1}^{I} d_i} \quad where \quad i = 1 \; to \; I \qquad (1)$$

(4) The technical objectives (TO), $t_j, \quad where \quad j = 1 \; to \; J$, are determined by the engineers and are set to be the measurement of the standards to which customer requirements are satisfied.

(5) The relationship between CCR and TO is denoted by R_{ij}. The engineers decide this value based on their skill, best practice, knowledge and experience. When there is no relation between CCR and TO then R_{ij} is assumed to be 0 and when there are relations then values like 9, 7, 5, 3, 1 are used.

(6) The weight of the TO is denoted as $w_t(t_j)$ and defined as:

$$w_t(t_j) = \frac{(\sum_{i=1}^{I} w_r(r_i) R_{ij})}{(\sum_{i=1}^{I}\sum_{j=1}^{J} w_r(r_i) R_{ij})} \qquad (2)$$

$$where \quad i = 1\ to\ I \quad and\ j = 1\ to\ J$$

Assumptions and notations used in FMEA

FMEA reports produce the Risk Priority Numbers (RPN) for each of the technical objectives. Companies adopt different ways of measuring risks and calculating RPN, but this research has taken the standard and most used approach. Traditionally the RPN is calculated by:

RPN=Occurrence (O) × Severity (S) × Detection (D)

(1) The RPN of a technical objective $t_j, \quad where \quad j = 1\ to\ J$ is denoted by RPN_j.

(2) The weighted mean of a RPN for each TO is denoted by $w_{RPN}(t_j)$ as:

$$w_{RPN}(t_j) = \frac{RPN_j}{\sum_{j=1}^{J} RPN_j} \qquad (3)$$

Assumptions and notations used in TOTA

Testing activities, which measure the effectiveness of technical objectives, are denoted as $T_s \quad where \quad s = 1\ to\ S$. The relationships between technical objectives and testing activities are denoted by M_{js}. The values in the matrix are defined as 5, 3, 1, or 0. Similar convention of HoQ , i.e. the odd values are used.

Overall equation for weighting Testing Activities

The importance of a Testing Activity is denoted as $I_T(T_s)$ and defined by

$$I_T(T_s) = \sum_{j=1}^{J} w_t(t_j) * w_{RPN}(t_j) * M_{js} \qquad (4)$$

Finally the weight of a Testing Activity is normalised using

$$w_T(T_s) = \frac{I_T(T_s)}{\sum_{s=1}^{S} I_T(T_s)} \qquad (5)$$

Tests						Importance of Test	Importance of Test normalised
Heavy duty endurance cycle	5	5	5	3	1	0.78	0.2
Thermal cycle	5		5			0.53	0.14
Emission measurement		5		5	5	0.48	0.12
Transient machine cycle	5	3	3	5	3	0.84	0.22
Virtual thermal mechanical analysis	5		5			[illegible]	0.14
Virtual fatigue analysis	5		3			0.46	0.12
Noise and vibration measurement	3			1		0.26	0.07
Relative criticality number weight	0.22	0.11	0.22	0.22	0.24		
RPN from FMEA	170	90	170	170	190		
Technical Objectives / **Customer requirements**	Maximum cylinder pressure	Fuel consumption	In cylinder temperature	After treatment size	Engine out emissions	**Customer Importance**	**Customer Requirement Weight**
Operating cost	5	9	1	1	5	5	0.14
Exhaust emission	5	5	5	9	9	5	0.14
Power vs Speed	9	5	9	9	5	4	0.11
Sound power	9	1	1	5	1	5	0.14
Sound quality	9	1	1	5	1	4	0.11
Transient response	9	9	5	5	5	[illegible]	0.14
Durability (B10 life)	9	1	9	1	1	4	[illegible]
Cost	9	1	5	9	5	5	0.14
Technical Characteristics weight	0.31	0.16	0.15	0.21	0.16		

$$w_T(T_s) = \frac{I_T(T_s)}{\sum_{s=1}^{S} I_T(T_s)}$$

M_{js}

$$w_{RPN}(t_j) = \frac{RPN_j}{\sum_{j=1}^{J} RPN_j}$$

$$w_r(r_i) = \frac{d_i}{\sum_{i=1}^{I} d_i}$$

R_{ij}

$$w_t(t_j) = \frac{(\sum_{i=1}^{I} w_r(r_i) R_{ij})}{(\sum_{i=1}^{I} \sum_{j=1}^{J} w_r(r_i) R_{ij})}$$

Figure 6.4 Equations for analysing the TPT Tool

Sequence of performing calculations

The steps in which these equations should be used is given below:

1: Calculate the weight of technical objectives (TOs) using equation (2).

2: Calculate the relative weight of RPNs of TOs using equation (3).

3: List the tests assigned with each of the TOs.

4: Rank the relationship between a Testing Activity and a TO.

5: Calculate the importance of a Testing Activity using equation (4).

6: Normalise the weight of a Testing Activity using Equation (5).

6.2.3 The method of populating the tool and analysis

The method of populating the tool and prioritising testing activities is a seven step process, which is presented in Figure 6.5.

The details of each of the seven steps in Figure 6.5 are as follows:

1. Selecting the product should be dependent on the intended use of the method and the time of the product development process, i.e. the stage of the product development process, because the focus of testing changes at different development stages. An analysis for a particular product will consider a specific set of testing activities, but a broader list is needed to accompany product variants and diverse operating environments. This process prioritisation should start between Gateway 1 to Gateway 2 along with the QFDs and FMEAs, but needs to be executed after QFD and FMEA analysis are completed. However, this should be a continuous process, constantly updated as the values in QFDs and FMEAs change.

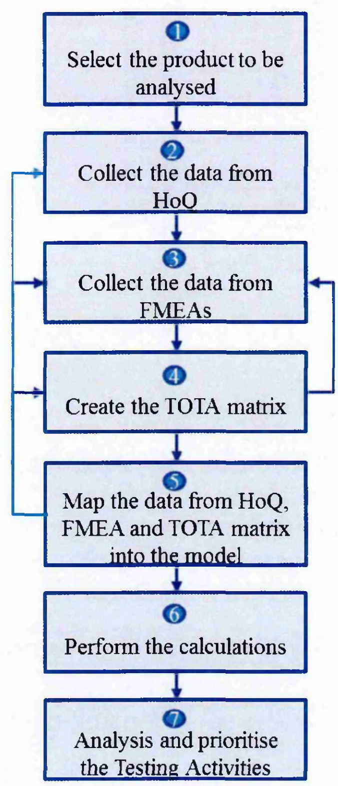


Figure 6.5 Overview of the Testing Activities prioritising method

2. Again, based on the intended use, a product is selected. The appropriate set of data (i.e. CCRs, TOs and relative weight of TOs) should be collected from HoQ. A general set of CCRs will include the environment and use conditions, while a baseline set will only consider product dependent CCRs.

3. The data collected from FMEAs should be based on the product selected in step 1 and the TOs selected in step 2. It is assumed that each of the CCRs will be satisfied with one or more TOs. These TOs will be analysed in FMEAs. Also, the risk priority number (RPN) of each TO is collected from FMEA. However, the

value of an RPN changes as the product development process progresses. At the beginning of a project, the risk of failing a TO can be assumed to be high but as the development progresses this risk is minimised through design changes or as preventive action is taken. Therefore, as the RPN value of a TO changes, the results of this method will change accordingly (as the back arrows show). Hence, this method should be adjusted each time the FMEA is altered.

4. The TOTA matrix should be created and filled based on the TOs and related Testing Activities collected from FMEAs. This is a critical step. Each of the TOs needs to be verified or validated through one or more Testing Activities, which is assessed in FMEA and can readily be used. In this step, the focus is on identifying and mapping TOs with Testing Activities.

5. To map the data from steps 2, 3 and 4, a matrix has been proposed in Figure 6.3, called a Test Prioritising Tool (TPT). The purpose of this mapping is to visualise the relationships between CCRs and TOs and between TOs and Testing Activities. In this arrangement it is easy to follow which Testing Activity is performed to evaluate a TO such that a related CCRs is met by achieving that TO.

6. The next step is to sequentially perform the calculations using the equations (1), (2), (3), (4) and (5). As mentioned before, equation (1) and (2) are standard equations of calculating weights of TOs (adopted from elsewhere). They can be applied if a company does not have a way of calculating the weight of TOs. Otherwise, the values of the weight of TOs can be used directly from HoO.

7. The calculation in step 6 produces the weight of each of the Testing Activities. This provides a measurement to rank these activities, providing that the higher weight means more important Testing Activity. Based on this ranking these Testing Activities can be prioritised.

Back arrows means that when any information is changed in HoQ and/or FMEA, the values and the calculations of the TOTA matrix are required to be adjusted accordingly.

In Chapter 4 (Section 4.3) the tools that are used in the company for testing planning were discussed. A schematic of data flow between these tools is presented in Figure 6.6. The TPT tool is introduced in that flow diagram. A representation of the TPT tool in the flow diagram and flow between HoQ, FMEA and TPT is also shown in Figure 6.6. It shows the decompositions of product levels in which this analysis can be performed.

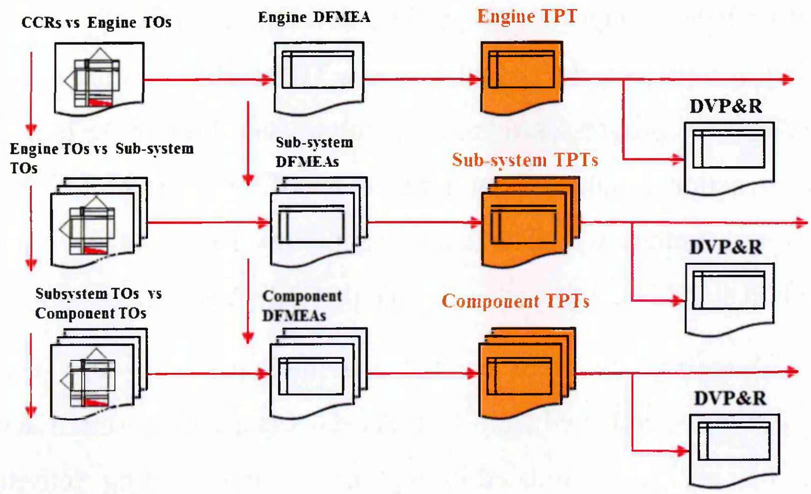


Figure 6.6 A schematic of the TPT tool into the process of testing planning and the data flow between these tools

6.3 Illustration of the method with an example

In this section, an example is presented to demonstrate the method proposed in Section 6.2.3. This example has been created, and is not based on real data. This example was performed on level 1; the Engine level.

The aim of the exercise was to deliver the initial idea of the method to the engineers. The engineers were also asked to evaluate the method based on the experiences they have gained through the exercise. This example was implemented with two engineers, Engineer 3 and Engineer 8, who were involved in the interview process (Table 3.1). Engineer 3 is the validation team leader who has good knowledge in designing the validation programme, identifying critical Testing Activities and planning the testing activities. Engineer 8 is the core Engine Mechanical System Team Leader, who recently has been given the challenge to make QFD and FMEA work more efficiently and shift these tasks towards the front of the gateway stages. Engineer 3 was the key person who actively implemented the process and was present during the whole (2 hour) period of the exercise. Engineer 8 was present for the first hour. This example (shown in Figure 6.7) was kept relatively simple to make it easily understandable without detailed technical expertise and specific domain knowledge. No particular product was in mind, rather a set of general customer requirements that are applicable to most of the current engine programme was considered.

The method has seven steps as shown in Figure 6.5; only the first 5 steps were performed during the exercise session. The author performed step 6, next day on her working space. However, all the equations suggested in this method were presented to Engineer 3 and

Engineer 8. They were aware of the whole process and the final results were sent back to Engineer 3 by email. The following sections trail the steps of this exercise.

1. Select the product and the level of the product to be considered

No particular product was in focus for this exercise. Selection of a particular product may depend on a number of criteria: by application, power range, emission standards; also by configuration – number of cylinders, aspiration, mechanical or electronic.

2. Collect the data from HoQ

As mentioned before, this is not a real example, and it was created mostly by one engineer. Data was not collected from an existing HoQ and the HoQ matrix was created during the exercise following the standard method. The critical customer requirements (CCRs) selected are generic requirements, which can be applicable across current engine development programmes in the company. These are requirements that frequently came into their discussions because they are critical. Engineer 3 could recall them readily during the exercise. Engineer 3 ranked these CCRs based on his judgement and understanding. A list of Technical Objectives (TOs) was generated according to these CCRs. Engineer 3 qualitatively ranked the relation between CCRs and TOs initially by putting H, M, L or keeping the cell empty when there is no relation. Later these letters were replaced with 9, 5, 1 or 0 respectively.

3. Collect the data from FMEA

In this step RPNs associated with each TOs should be collected from FMEAs, but again no data was collected from a real FMEA. Engineer 3 ranked the TOs based on the risk of not meeting these. Engineers were hesitant to put any real number but expressed with H, M, L, which meant high risk, medium risk and low risk respectively. Later, the author translated these latters into numbers (shown in Figure 6.7).

4. Create the TOTA matrix

Engineer 3 listed some of the engine level Testing Activities, which are usually performed on an engine during a validation programme to ensure these TOs. He mapped the relationships between TOs and Testing Activities. Again, the relationship between TOs and Testing Activities were expressed by H, M, L and N, which were replaced latter with 9, 5, 1 and 0 respectively.

5. Map the data from HoQ, FMEA and TOTA matrix into the model

As the author previously planned to organise the data into the model presented in Figure 6.5, a blank template was provided to the Engineer 3 and Engineer 8, at the start of the exercise. Engineer 3 was writing on boxes on the template in each of previous steps 1-4. So there was no need to map the data from HoQ, FMEA and TOTA matrix, as the data were already organised on the template.

Tests						Importance of Test	Importance of Test normalised
Heavy duty endurance cycle	5	5	5	3	1	0.78	0.2
Thermal cycle	5		5			0.53	0.14
Emission measurement		5		5	5	0.48	0.12
Transient machine cycle	5	3	3	5	3	0.84	0.22
Virtual thermal mechanical analysis	5		5			0.53	0.14
Virtual fatigue analysis	5		3			0.46	0.12
Noise and vibration measurement	3			1		0.26	0.07

Technical Objectives / Customer requirements	Maximum cylinder pressure	Fuel consumption	In cylinder temperature	After treatment size	Engine out emissions	Customer Importance	Customer Requirement Weight
Relative criticality number weight	0.22	0.11	0.22	0.22	0.24		
RPN from FMEA	170	90	170	170	190		
Operating cost	5	9	1	1	5	5	0.14
Exhaust emission	5	5	5	9	9	5	0.14
Power vs Speed	9	5	9	9	5	4	0.11
Sound power	9	1	1	5	1	5	0.14
Sound quality	9	1	1	5	1	4	0.11
Transient response	9	9	5	5	5	5	0.14
Durability (B10 life)	9	1	9	1	1	4	0.11
Cost	9	1	5	9	5	5	0.14
Technical Characteristics weight	0.31	0.16	0.15	0.21	0.16		

Figure 6.7 An example of the prioritising method

6. Perform the calculations

The author transferred the data to an Excel spreadsheet and performed the calculations suggested in equations 1 to 5. The results are shown in Figure 6.7.

7. Analysis and Prioritise the Testing Activities

The right column of the top matric in Figure 6.7 shows the results. From these results, it is identified that 'transient machine cycle' has the top score (i.e. 0.22) and 'noise and vibration measurements' has lowest score (i.e. 0.07) among these tests. This high score of 'transient machine cycle' indicates that this test could be the most important among these tests. But the engineer's judgement still played an important role in evaluating these scores. For example, both 'thermal cycle' and 'virtual thermal mechanical analysis' scored 0.14, which means both of these tests are equality important. However, engineers mentioned that 'thermal cycle' is a crucial test for validation that are performed in Design Verification (DV) and Product Validation (PV) phases, whereas 'virtual thermal mechanical analysis' is essential before an actual 'thermal cycle' test can be performed. 'Virtual thermal mechanical analysis' is performed in System Demonstration (SD) and Design Verification (DV) phases, providing the inputs and boundary conditions for 'thermal cycle' test measurement.

Findings from the exercise

This exercise provided a few important insights. Engineer 3 found it difficult to fill the QFD part. He was quite confident to determine the tests that will be needed for system level validation tests but it was not easy to determine the initial customer requirements that they were testing for. This indicated that there was a communication gap or information gap between QFD and FMEA teams. Engineer 3 found this to be a useful exercise and thought this approach could be useful for understanding the overall system in an effective way. However, he also thought this approach might be too difficult to implement in detail when there are hundreds of failure modes need to be considered.

6.4 Evaluation of the method

The method was evaluated on two measures; firstly, ease of use and secondly, added benefit of using this method against a current method.

'Ease of use' was measured by considering the time and effort it takes to complete the process and how easily it can be learned. According to Engineer 3, this is a simple process which can be learned effortlessly but the effort and time required to implement the method depends on how easily and readily the results can obtained from HoQs and FMEAs. Engineer 3 also mentioned that this process may be easy to implement in engine and sub-system levels but could be difficult to implement in detailed component levels because, there can be many failure modes in component level and not all of these modes can be easily mapped to the technical characteristics or overall customer requirements.

Engineer 3 also found the TOTA matrix is useful and provide added benefit because it captures the relationships between tests and associated TOs. Although, in an FMEA, these relations are presented, this way of is particularly useful as can be used to visualise how a test is connected to multiple TOs, or multiple TOs are connected to a test. In this presentation, the relation between Testing Activities to TOs and TOs to CCRs are clearly demonstrated, therefore it is visualised spontaneously which test is performed to ensure a TO such that a CCR is met. This approach was also discussed with Engineer 1 in a later meeting and he commented,

> "this is a much more structured approach to how you should go about developing an engine logically rather than randomly".

6.5 Summary

A method of integrating QFD and FMEA for prioritizing testing activities has been proposed this chapter, which can improve the current test prioritisation process of the company.

Verification and Validation managers require good understanding of both technical and customer requirement for effective test planning. Currently, engineers prioritise testing activities based on their experience, focused on most recent and major issues. Although engineers use their learning and educated judgments, RPN is the only measurement that is currently used in the case study company to prioritise the testing activities. The proposed method brings other factors, i.e., voice of customers and relative importance of the testing activities into this measurement.

The proposed method also provides an improved process in the sense that testing activities can be focused on critical aspects of both technical and customer requirements. However, this method is limited in one specific aspect. This study has found that there can be many TOs in FMEAs, which are defined internally by engineers that might not directly satisfy CCRs. These TOs are essential for many reasons (i.e. internal reasons, explained in Section 4.4.2). Although, these auxiliary TOs are critical and are required to be dealt with, these TOs cannot be considered directly in this method because these TOs cannot be directly mapped to the TOs in HoQ. Further research will be required to address this issue.

Chapter 7 A DSM based modelling for organising testing activities

The empirical study has found that the planning of testing activities is complex and can be difficult to manage because it depends on many factors such as, information coming from design, resources availability and appropriate timing. Also, the interrelations and learning between the testing activities are critical for the planning of testing activities. If dependencies and interrelationships are not known and captured appropriately, it is difficult to maintain and monitor the information flow between activities; therefore, information can be lost.

In addition the company is continuously trying to improve and optimise its testing processes by aggregating, merging, or disaggregating testing activities, where appropriate. Rising fuel costs and reduced development time provide an additional pressure to reduce the number of tests the company is currently running. Engineers, working on a particular product over many years acquire high levels of knowledge and can continue this optimisation. A large amount of information is implicit and is not expressed explicitly in reports and documents but is part of the engineers' tacit knowledge. It is important to extract the undocumented knowledge from the engineers and to capture it in an appropriate form.

In the case study company, dependencies between tests and components are maintained and managed as lists in Excel spreadsheets. It is difficult to visualise the dependencies and interlinks in this form. For the purpose of representing the dependencies and interrelations that exist between components and testing of a product, a Dependency Structure Matrices (DSM) based modelling approach is proposed in this chapter.

Section 7.1 outlines a brief introduction to product and process modelling; Section 7.2 describes the proposed tool for capturing the dependencies. An example is used in Section 7.3 to demonstrate the tool. Section 7.4 provides a brief summary.

7.1 A brief introduction to product and process modelling

The modelling of aspects of design products and processes is a large subject area and will only be addressed briefly in this section. A thorough review of models of designing is given by Wynn & Clarkson (2005) and Browning et al. (2006) who reviewed key concepts in process modelling.

According to Suh (2001), the world of design is made up of four domains: the customer domain, the functional domain, the physical domain and the process domain. In the case of product design, customer domain consists of the needs and requirements that the customer wants from the product. The functional domain consists of the functional requirements, also defined as specification and constraints. The physical domain consists of the design parameters that are chosen to deliver the functional requirements. And finally, process domain specifies the process that can produce the design parameters. These four domains are logically connected and represent the design domains. Also, there is a large amount of connectivity between the products that a company produces and the processes that it uses (Earl et al. 2000). In this chapter only physical (product) and process domains are considered.

Modelling represents the real world in a model for a specific purpose. The development of an accurate and appropriate model depends on the understanding of the architecture of the elements (often called organisation of the elements) of the product or activities of the process. *Product architecture* refers to the components and their interactions within physical artefacts and process *architecture* refers to the activities and the interaction that accomplish the desired work, such as design and testing of a product; these definitions are adopted from (Eppinger & Browning 2012).

7.1.1 Tools for modelling

Different tools are used for modelling the product/process including a directed graph approach (Lawler 2001), Program Evaluation and Review Technique (PERT), Graphical Evaluation and Review Technique (GERT), Critical Path Method (CPM), and Dependency Structure Matrix (DSM). The latter is often referred to as a Design Structure Matrix when used in the design context. Among these tools, the DSM representation has been used successfully for project management (Yassine et al. 2008), information management in product development (Shamsuzzoha & Helo 2011) and concurrent engineering (Chen & Lin 2003, Yassine & Braha 2003). This study uses DSM for modelling the dependency of product and process domains. A DSM is particularly useful because of its key strengths in highlighting iterative loops (Eppinger 1994), by representing nodes in both lower and upper triangular parts of the DSM matrix. A DSM can be analysed to minimise feedback loops, i.e., trying to order activities to upper triangularise the matrix. Further advantages of DSMs have been described by other authors such as (Eppinger & Browning 2012, Yassine & Braha 2003, Lindemann 2012, Browning 2001). These include indication of relationship intensity, probability, or other attributes, which highlight important patterns in the design.

7.1.2 Dependency Structure Matrix (DSM)

The Dependency Structure Matrix (DSM) is an established method for capturing the complex interaction of design tasks (Browning 2001, Steward 1981) and is widely used for capturing the dependencies. A DSM provides a simple visualisation of the dependencies between elements of a system. It is a square matrix (as shown in Figure 7.1), where the rows and columns are named and ordered identically. The elements of the domain of analysis are assigned along rows and corresponding columns. An off-diagonal mark indicates the dependency of one element on another.

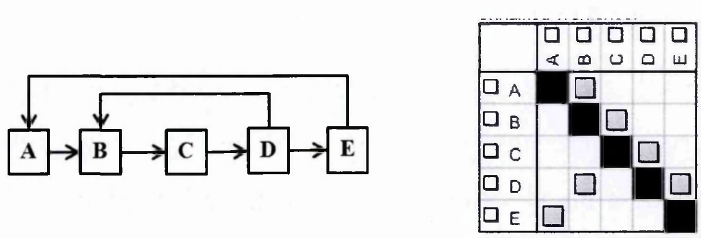


Figure 7.1 A representation of Dependency structure matrix (DSM)

There are two conventions for reading DSMs. This thesis uses the clockwise IC/FBD convention and details can be found in Lindemann (2012). Reading down a column reveals the input sources, while reading across a row indicates output links. Thus, in Figure 7.1, element B provides an influence on element C and it depends on something from elements D and A. Diagonal marks do not represent any relation, because a relation from an element to the same element is not permissible in DSMs.

There are two main types of DSMs: static and time-based. Static DSMs represent component-based or team-based system elements. A component-based DSM (CDSM) represents interactions between components in complex system architecture. A team-based DSM (TDSM) represents interactions between people and/or groups within an organisation. In time-based DSMs, such as activity-based and parameter based DSMs, the ordering of the rows and columns indicates a flow through time. In an activity-based

DSM (ADSM), the order of activities in the rows and columns indicate some appropriate 'time-ordering', possibly in terms of starting time. Any feedback marks will appear below the diagonal. A parameter-based DSM represents the relationship between design decisions and parameters.

When a DSM represents multiple domains then it is called a Domain Mapping Matrix (DMM). It allows linking between two domains (one domain is in rows and other is presented in column) to analysis the relationships. DMM can be square or rectangular but unlike DSMs the diagonal marks of a square DMM represent a relation between the two domains of modelling. The convention of reading a DMM is same as the DSM.

DSMs and DMMs are used to model whole systems consisting of multiple domains; each can have multiple elements, connected by various relationship types. This model is called a Multiple Domain Matrix (MDM). MDM allows analysing a system's structure across multiple domains into a single matrix. Details on MDM can be found in (Maurer & Lindemann 2008).

7.1.3 Method of analysing matrices

A number of algorithms have been developed to analyse and manipulate the DSMs. Static DSMs are usually analysed with clustering algorithms. Time-based DSMs are typically analysed using sequencing algorithms. The goal of clustering is to find subsets of DSM elements i.e. finding the groups of elements that are interconnected among themselves, while being little connected to the rest of the system. Sequencing is the reordering of the DSM rows and columns such that the new arrangement contains as few feedback marks as possible to the lower diagonal. There are many sequencing algorithms to minimise iterations. These can be found in (Eppinger & Browning 2012, Browning 2001, Eppinger et al. 1992).

However, none of these techniques were used in this research. The intention of this research is to capture and represent the current state of the product and process architecture in an efficient way and not to influence the structure of the current process. Also, since engines are such a mature product, with the company having many years of experience in designing the product, there is a little that can be improved using these techniques. A new technique for modelling components and tests is developed here.

7.2 A tool for modelling components and tests

Several underlying factors were recognised in the dependencies between tests:

(i) a change in one test might affect other tests; for example tests can render each other obsolete or invalid,

(ii) changes (due to failure, for instance) to a single component can affect testing plans, and

(iii) a change to one component of the product may result in changes to other components and lead to associated changes in tests.

These factors imply not only the interrelationships between tests but also that connecting each test with physical parts/components is essential for effective testing plan.

As mentioned before MDM is used for modelling multiple domains. Therefore, a MDM is used for modelling:

(a) connections between components using a CDSM,

(b) relationships between components and tests using DMM and

(c) dependencies between tests using ADSM.

The modelling of the ADSM was the focus of this study. This study demonstrates a process of modelling an ADSM for the whole engine system considering different stages of the product development process. The other matrices, i.e. CDSM and DMM were modelled to link the testing activities to components, so that, the effect of changes in any domain to the whole system can be analysed.

CDSM The list and interactions of technical components of the product	**DMM** Relates the technical components that act on system activities
DMM Relates the testing activities that act on technical components	**ADSM** The list and interactions of testing activities of the product

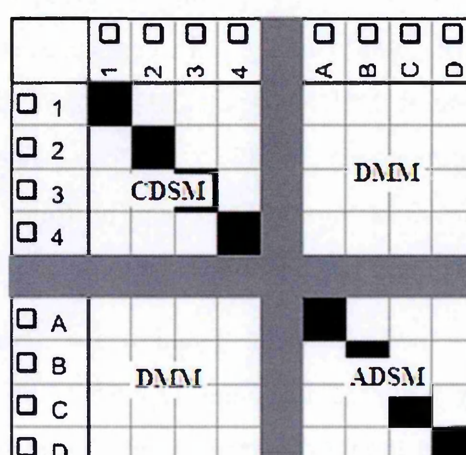


Figure 7.2 A schema of Multiple Domain Matrix (MDM)

This proposed MDM has four quadrants (shown in Figure 7.2). The Top-left quadrant was used for modelling the CDSM, bottom right quadrant for ADSM. Both bottom-left and top-right quadrants can be used for modelling the DMM. According to Bartolomei et al. (2012) bottom-left DMM relate the activities that act on technical components and top-right DMM relates the technical components that act on system activities (adopted from (Bartolomei et al. 2012)).

For the model proposed in this exercise, only the bottom-left DMM was used to relate the testing activities that act on components. Usually testing activities are performed on components. Any changes in a component can change the procedure, requirements and parameters of tests. Normally, in the case study company, a component does not change because a test procedure has changed. The relation of components that act on testing activities was not considered and was not modelled in this thesis. Therefore the top-right quadrant is left blank.

7.2.1 CDSM

CDSMs are used to represent the linkage between the components of the system. The term 'linkage' is adopted from (Jarratt 2004). Some interaction types are spatial, energy, material and information, although there can also be a wide range of other types of linkages which can connect two components of an engine, identified by (Jarratt 2004). This research will only consider geometric linkages, i.e. mechanical steady state and spatial. There is a mechanical steady state linkage between two components when components are in physical contact and they impose a steady state mechanical force on each other and spatial is established when components are touching or adjacency is important (Jarratt 2004). Both relations are symmetrical.

7.2.2 ADSM

ADSMs represent the information flow between the activities in a chronological order. There is a logical sequence that the case study company follows while organising testing activities. First, testing activities are organised as these will be performed during the stages of the product development process, i.e. first, System Demonstration (SD) tests, then Design Verification (DV) tests and finally Product Validation (PV) tests. At each stage, component level testing will be performed first, then subsystem level and then system level tests. Also at each level, performance testing is performed before mechanical testing.

Usually components are tested in the SD phase, sometimes in the DV phases and very rarely in the PV phase. Components will require major and costly changes after DV phase. Similarly, subsystems are tested in the SD and the DV phases, and very rarely in the PV phases. Engine level testing happens across the SD, DV and PV phases.

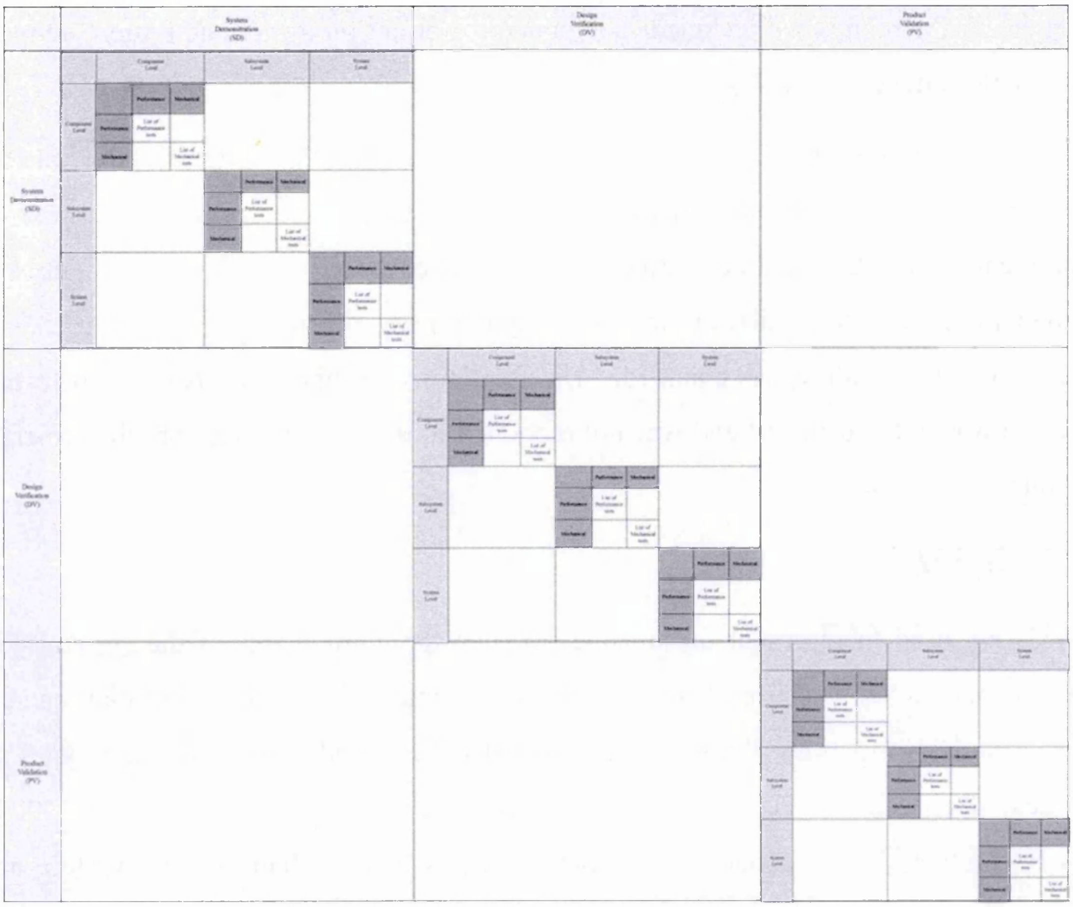


Figure 7.3 The schematic of overall ADSM

Knowing these rules, these following steps should be followed to capture the sequence of testing activities of the whole engine system during the product development process:

i) First, group testing activities based on stages, i.e. SD tests, DV tests and PV tests.
ii) Second, group each group into subgroups based on component level, subsystem level and system level tests.
iii) Third, each subgroup is further grouped into sub-sub-groups by separating performance tests from mechanical tests. This provides a list of performance tests and a list of mechanical tests.
iv) Fourth, find the sequence of these testing activities in each group created in third step.

The schematic of overall ADSM is shown in Figure 7.3 and the arrangement of the testing activities into the structure of the ADSM is shown diagrammatically in Figure 7.4.

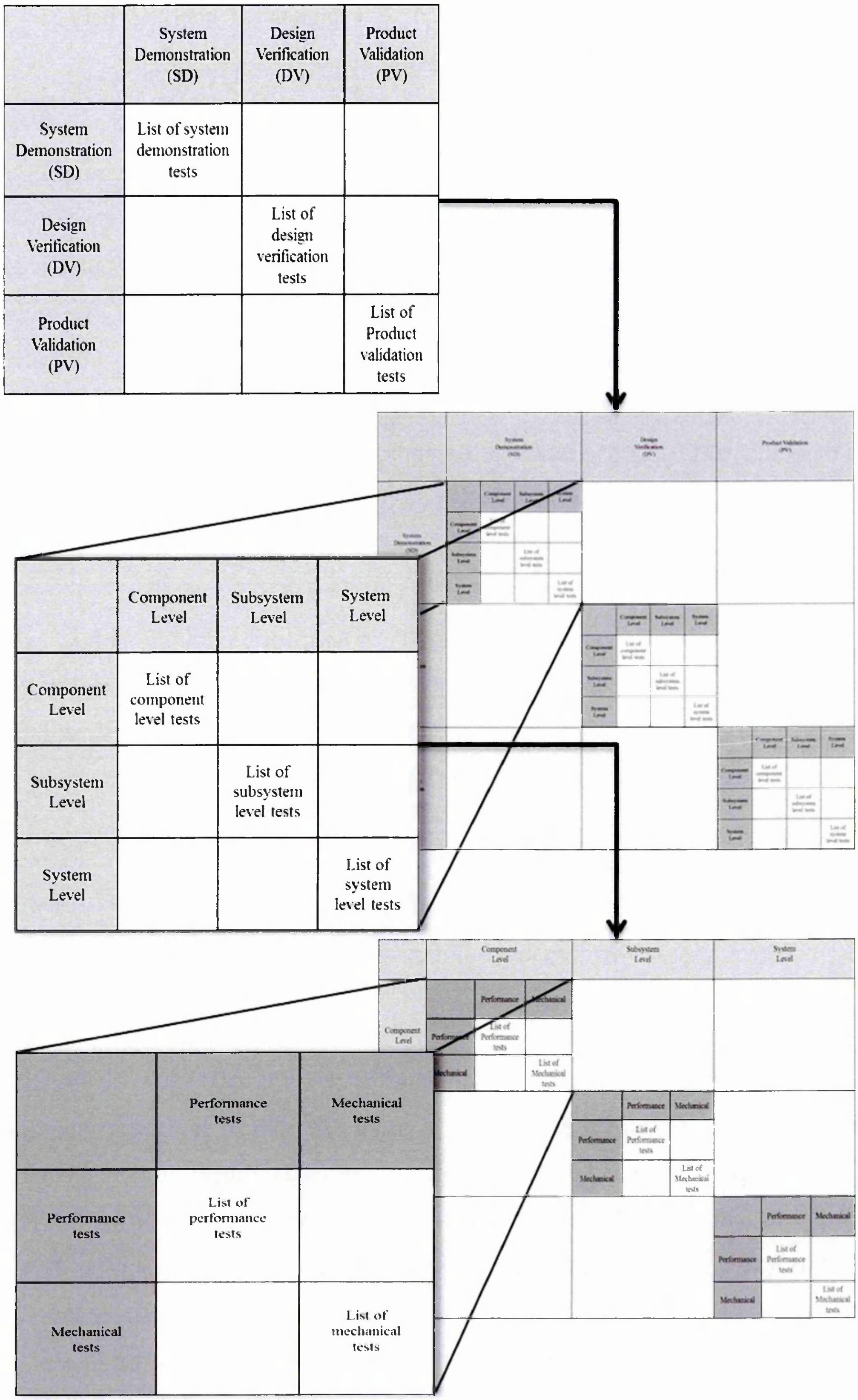


Figure 7.4 The arrangement of the testing activities into the structure of the ADSM

7.2.3 DMM

The Domain Mapping Matrix (DMM), which was presented in Figure 7.2, relates two domains. This DDM is essentially used to establish a mapping between tests and components, i.e. to model which component requires which tests.

It is important to note that, this DMM doesn't mark all the components or sub-systems in the case of a system level test; on the contrary it only marks the components that are the focus of the test. All the components can be subjected to a system level test but often a single component or a subsystem is the focus of a particular test. For instance, a 'gross thermal test' is performed on an engine to examine the thermal fatigue of a cylinder head and cylinder head gasket; therefore the focus of this test is the cylinder head and cylinder gasket. Although, other components of the engine will be subjected to this test, these are not modelled in this DMM initially. So, for the above example, gross thermal test has a relation only with cylinder head and cylinder gasket in this DMM.

7.3 An example MDM

This section illustrates the process that has been used for modelling with an MDM. First, there is a description of how to build the MDM. Second, findings were established and third, it is shown how to improve MDM modelling through gaining insights from the company. Finally, ways of analysing the MDM are discussed.

7.3.1 Building MDM

This exercise followed the standard method of building DSM (adopted from (Dong 2002)). The process of creating and building the MDM follows the steps:

7.3.1.1 Preparation

To collect appropriate system decomposition and the accuracy of the dependence relationships, it is recommended to execute the DSM by gathering a group of engineers/experts from different functional groups of the organisation and asking them to collectively list the different subsystems that comprise the system as a whole (Dong 2002). But it was not possible to gather engineers from different groups at the same time, because the company was going through a significant business change and most engineers were extremely busy during the period of this research. Engineers also could not assign the amount of time that was required for building a meaningful MDM. Thus it was decided to choose a core component (consisting of several parts) but small enough, so that any single engineer (involved with this research) has basic understanding of that component. Therefore, it wasn't necessary to bring all engineers together for creating and evaluating it. The interim work that follows on the next steps was completed by the author before placing this example to the engineers.

7.3.1.2 Define the system and its scope

An engine has hundreds of components and hundreds of associated tests, a small part of the engine, i.e. piston and connection rod were selected to model as an MDM, with the aim of understanding the organisation of testing tasks and components.

The piston and connecting rod are core components of the engine because the majority of the connections in the engine concern them. Also, the verification and validation of a piston is a complex process because it goes through many tests and it is tested many times at subsystem and system level. Figure 7.5 shows a decomposition of a piston and connecting rod.

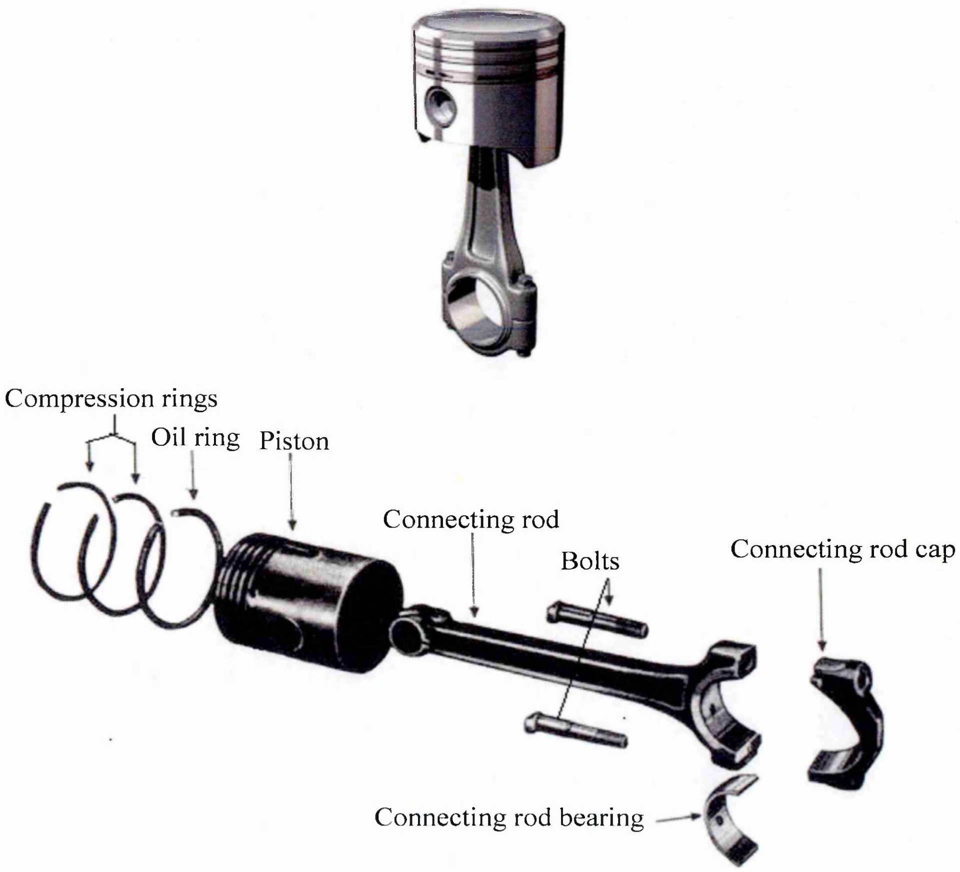


Figure 7.5 Decomposition of piston and connecting rod

7.3.1.3 List all the system elements

Ideally, the system elements should be collected from the design documents and engineers' suggestions. For this exercise, the ideal place was to start with a DFMEA (i.e.

Design FMEA). Usually, each component, each subsystem and the engine itself have a DFMEA, where testing activities for verification and validation purposes are listed. More than one DFMEA will need to be considered for a component to scan through all their associated tests because a component may subject to subsystem and engine level tests. Often DFMEAs are not created for small parts but embedded with a component or a subsystem, which themselves are subjected to tests.

Due to confidentiality reasons, it was not possible to get access to the actual DFMEAs. An initial set of design elements i.e. the list of components and the list of associated tests were defined based on a reference book called "Automotive Engine and Vehicle Technology" (Chisholm 1984). Initial lists of components and tests are shown below in Figure 7.6. These tests confined to design verification (DV) stage. But later found that some of these are performed in other phases of product development.

List of components   **(for CDSM)**
1. Piston
2. Compression rings
3. Oil ring
4. Connecting rod
5. Conn. rod bearing
6. Connecting rod cap
7. Bolts

List of Tests   **(for ADSM)**
a. Stress analysis
b. Strength test of piston head
c. Heat-expansion measurement
d. Performance
e. Temperature measurement
f. Wearing test
g. Compression testing
h. Fatigue resistance
i. Load carrying capacity
j. Vibration test

Figure 7.6 Lists of components and tests considered for MDM

7.3.1.4 Establishing the relations between the system elements

To establish the relations between system elements a meeting was arranged with Engineer 1. The whole process of establishing the relations was completed in a single meeting (148 minutes). Engineer 1 was familiar with the process of creating DSMs. He has created many DSMs, during his work at the company; also he was involved in the process of creating change prediction method (CPM) by another student (Jarratt 2004). Initially, the scope of this MDM and the linkages were explained to the Engineer 1 during the exercise. A template, created in Excel (shown in Figure 7.7) and printed on an A4 paper, was used for this exercise. The engineer was requested to think aloud, so author was also able to fill in the matrices. The audio of this meeting was recorded.

At the beginning of the exercise, Engineer 1 suggested that the list of tests should be updated, as some were currently not used in the company; also some of these had been renamed. For instance, compression testing was renamed as blow-by test. Engineer 1 has not been recently involved in verification and validation planning.

A single quadrant of the MDM was filled at a time. There was no suggestion from the author about which quadrant to fill in first. Engineer 1 filled the DMM first, by considering which test will be performed on which component. Second, the ADSM was completed. The arrangements of the tests have been significantly changed, as initially, these were not chronologically ordered. He followed the basic rules as explained in Section 7.2.2, i.e. arranging the performance testing before mechanical testing. Also component level tests were placed on the top rows, before system level tests.

	1	2	3	4	5	6	7		a	b	c	d	e	f	g	h	i	j
1. Piston																		
2. Compression rings																		
3. Oil ring																		
4. Connecting rod												Not considered						
5. Conn. rod bearing																		
6. Connecting rod cap																		
7. Bolts																		
a. Stress analysis																		
b. Strength test of piston head																		
c. Heat expansion measurement																		
d. Performance																		
e. Temperature measurement																		
f. Wearing test																		
g. Compression testing																		
h. Fatigue resistance																		
i. Load carrying capacity																		
j. Vibration test																		

Figure 7.7 A template for MDM

The CDSM was difficult to fill in, because as an engine manufacturer the company do not test each of these components individually. Mostly these components are tested whilst embedded in subsystems in various test procedures. There were some assumptions made during the completion of this CDSM, which arose from it being an ideal case.

The linkages were modified several times during the meeting, as, through the course of the exercise, Engineer 1 understood the linkage type and the purpose of this modelling. The sequence for filling the DMM, ADSM and CDSM were followed roughly, but Engineer 1 moved between the quadrants many times and altered the linkages. This indicated that these closely interlinked domains are better represented in a single matrix and Engineer1 could get a better picture of the whole system and uncovered many links.

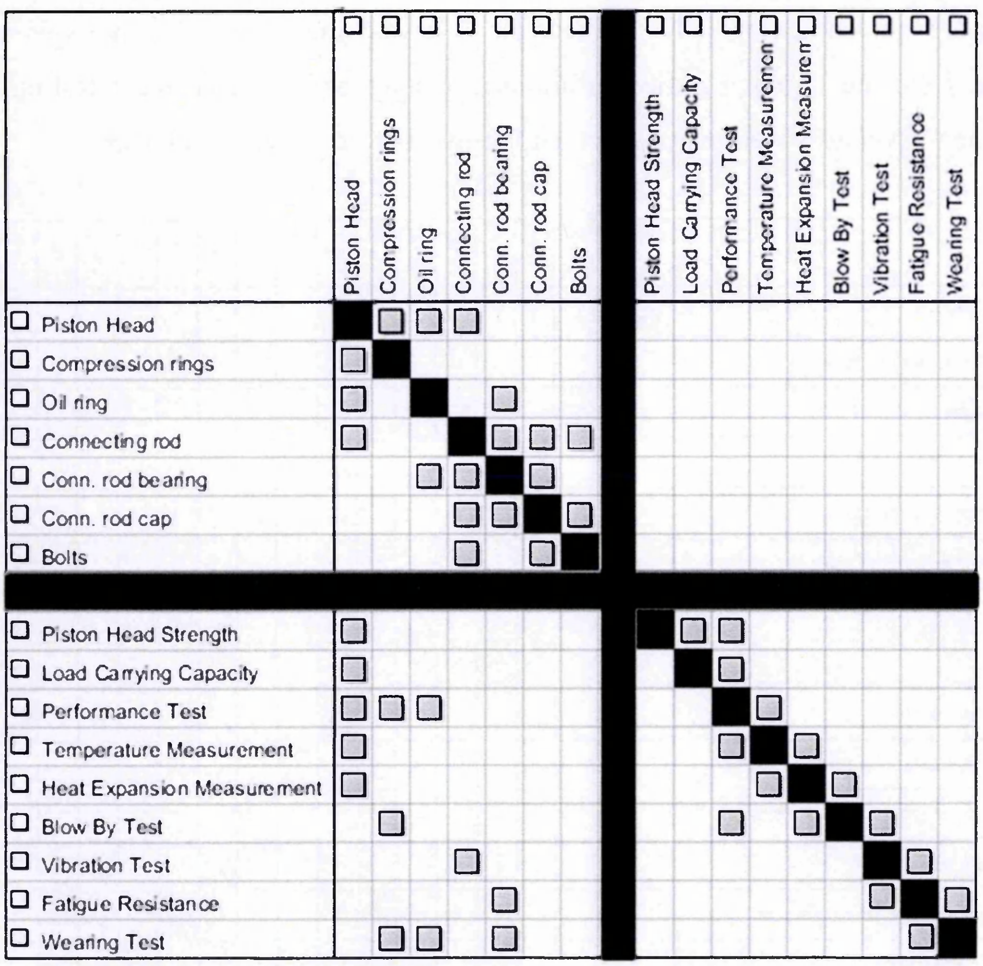


Figure 7.8 CDSM (top left quadrant), ADSM (lower right quadrant) and DMM (Lower left quadrant) for piston and connecting rod.

Subsequently, these DSMs were completed using the Cambridge Advanced Modeller (CAM) tool (Wynn 2011). Figure 7.8 shows the completed MDM that consists of a CDSM (top-left quadrant), a DMM (bottom-left quadrant) and an ADSM (bottom-right quadrant). The static CDSM captured the connections between piston, compression rings, connecting rod etc. are modelled in this matrix. As mentioned before, these relations are mainly mechanical steady state and spatial. For example, the oil ring may not have any spatial relation with the connecting rod bearing but oil passes through the bearing to the connecting rod to the oil ring to lubricate the piston and this process establishes a connection between oil ring and connecting rod bearing. The bottom left DMM shows the relationship between tests and components. It presents which tests are performed on which components. For example, the 'blow by' test is performed on compression rings and 'performance test' is performed on piston head, compression rings and oil rings. The ADSM captured the sequences and the information flow between different tests. For example, the 'blow by' test needs information from measurements of heat expansion.

7.3.2 Findings from the exercise

Initially, there were different views about this exercise from different groups of people. Engineer 1, who performed the exercise, considered it to be of high value from which the company benefits. On the other hand, Engineer 2 (who is working in this industry for last 15 years), was the verification and validation manager at the time of this study, with an understanding of the whole product development process commented

> "an engine is a very very complex thing with hundreds of parts and thousands of CAE/test activities. It will therefore not be possible to do this at an engine level. Even a single component can be complex".

This exercise looked rather difficult to the engineers (Engineer 2) who had not taken part in DSM creating activities previously. He also believed it is very time consuming and required a significant amount of effort.

It was noted in the course of the case study that there are ideal processes (engineers think what they should do) and actual processes (what they really do). During the testing planning, testing activities are organised in an ideal way (how it should be), but when a disturbance happen, an engine breaks during a test, for instance, then the company takes a 'fire-fighting' approach. That is it allocates resources or efforts into this unforeseen situation that must be dealt with immediately. These 'fire-fighting' situations might differ programme to programme. Hence, actual processes often deviate from the ideal plans that have been set down as plans at the beginning of the projects. When they were asked to explain their process, engineers frequently described the ideal processes, rather than their actual processes, unless they wanted to emphasis a particular situation or event.

In the company, tests are performed sequentially for two reasons:

(i) *information dependency*, i.e. one test needs information from another. For instance, a test Y needs some data or information from X, so X needs to be performed before Y.

(ii) *resource scarcity*, i.e., although X and Y may not have information dependency, both need the same resources that can only be used by one, creating a queue.

Engineer 1 did not explicitly mentioned about these differences in the links between testing activities during the exercise. Engineers assume these are known to others and do not explicitly explain the reasons behind the specific nature of a link between activities. If the reasons of putting a link are not explained explicitly, it limits the use of these matrices by others. Often this is why these matrices may not be transferable. They are only understood in detail by the modeller themselves.

During the exercise, Engineer 1 also provided a few important suggestions to improve the quality of the DSMs. As there are three matrices that are need to fill in this MDM, this is not a trivial task. The number of linkage increases as the number of components and testing activities increases. Starting with a manageable number of components and testing activities would be convenient, and can be expanded slowly as required. Also to simplify this fact and make it useable, it was recommended to model a single linkage in each DSM and define the linkage explicitly.

7.3.3 Improved MDM

The MDM model was corroborated by taking it back to the engineer, who has provided the information from which it was constructed. Later interviews with engineers provided some ideas for further improvement of this initial MDM.

As mentioned before, engineers continuously try to integrate the testing activities. In the case study company they create two lists: 'major' and 'minor' lists of testing activities. Core components are tested in these major tests and non-core components are tested through minor tests. More information about core and non-core component can be found in (Jarratt 2004). They aim to aggregate minor tests with major tests, where possible. For example, they can check the durability of an alternator bracket by installing it on an engine during the engine's mechanical test. Engineer 1 described this as 'aggregation'. However, is serves to take care here with the use of terms. This use of the term aggregation is different from that is used in DSM community. Aggregation, in the case study company's sense, can happen in two ways:

Test procedure merging: Two test procedures are merged into one test procedure and performed on an engine.

Parallel test procedures: Two different test procedures are performed at the same time on one engine. In this case the focus of the test can be different components.

As the company has many years of experiences of building and testing engines, the company's test procedures are highly advanced and tests are already integrated into single test procedure where possible. Also, the company is required to follow the standard testing procedures provided by the owning company and therefore, has little flexibility in changing them. However, they can organise and perform tests together without affecting the test procedures, depending on the number of test procedures that can be put into one test set up.

This study has found the aggregation, as a parallel procedure, of two tests often means that a minor test is partnered with a major test. Both tests are performed at the same time on a same engine in the test bed, without affecting each other. This study also identified four conditions that should be carefully monitored in this kind of aggregation:

(i) The tests are not actually on the same component (then it becomes 'procedure merging')
(ii) There might be no information flow between these tests (no sequential dependency)
(iii) Minor tests should not create interference (product interference, such as, vibration, noise, unintentional contact)
(iv) The time required for minor test should not affect the timing of major test.

For this study, the 'major test' and 'minor test' referred to in the company were renamed as *primary test* and *auxiliary test* respectively, to reflect these observations. The following definitions are proposed:

> *Primary test:* when a component is the focus of a test, then that test is the major test for that component.
>
> *Auxiliary test:* when a component is subjected to a test but is not the focus of that test then that test is a minor test for that component.

Engineer 1 remarked on the linking between primary and auxiliary test

> "if we have a matrix that relates tests saying that if you did a thermal test then you also get this test free that would be helpful. By looking at that you could say, well we want to do these 40 tests, but these 20 are actually unique, and we get other 20 free".

The DMM matrix was used for this purpose to represent the aggregation (as Engineer 1 termed it) of testing activities by clearly indicating the primary and auxiliary tests of a component. To do this, a link of a component with its primary test is marked with "P" and auxiliary tests are marked with "A".

In ADSM, two types of sequential links (discussed above) can be clearly indicated by marking "I" –when information flow, "Q"-when queuing. The priorities of testing activities play an important role when queuing required by arranging the highest priority activities first. Experienced engineers who were working on the particular product have a good source of knowledge and will be able to do prioritisation. But as mentioned before there is no methodological approach of calculating this and often can be subjective and based on the current "fire-fighting" situation. In the previous chapter a method of prioritising was proposed based on the relative weight of the testing activities, which could be a practical approach and adaptive to this situation. It was decided to weight testing activities, according to priorities, which can help during these analyses.

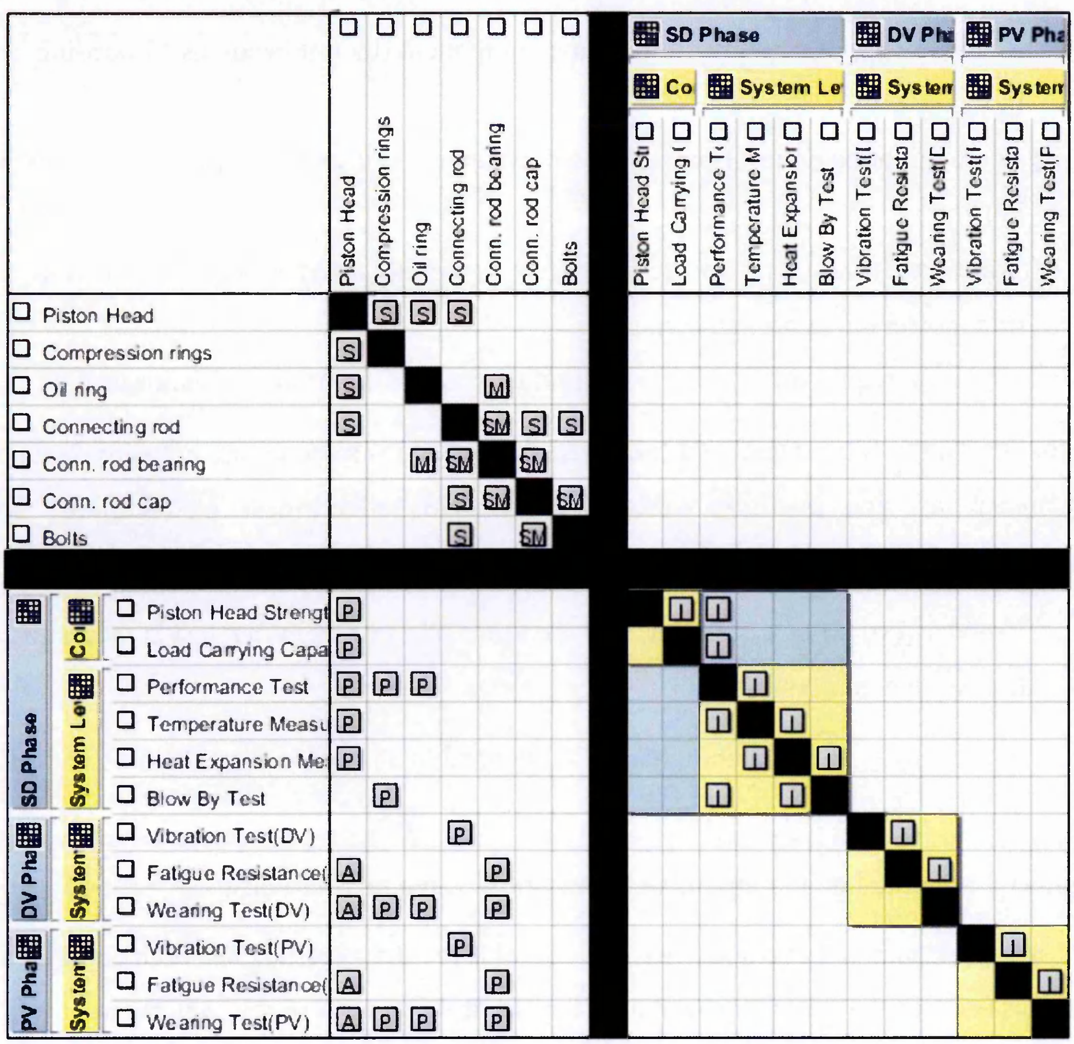


Figure 7.9. Final MDM for piston and connecting rod

In Figure 7.9, the ADSM in the bottom right quadrant has marks below the diagonal, which represent the iteration of the testing activities. From the ADSM matrix, it appears as if the vibration test and fatigue resistance are coupled. However, in reality, this represents the iterations of these tests in different phases. For instance, mechanical tests, like vibration test, fatigue resistance and wear test often require repeating during product validation (PV) phase when there is a hardware or production variation. Therefore it was planned to represent these in the matrix by explicitly mentioning the phase of these testing activities.

7.3.4 Method of analysis

The primary objective of this MDM modelling was to capture the dependencies between components and testing activities. The model succeeded in visualising the dependencies between these components and testing activities. This MDM can be used more widely for analysing how change propagates and what its effect on product development might be.

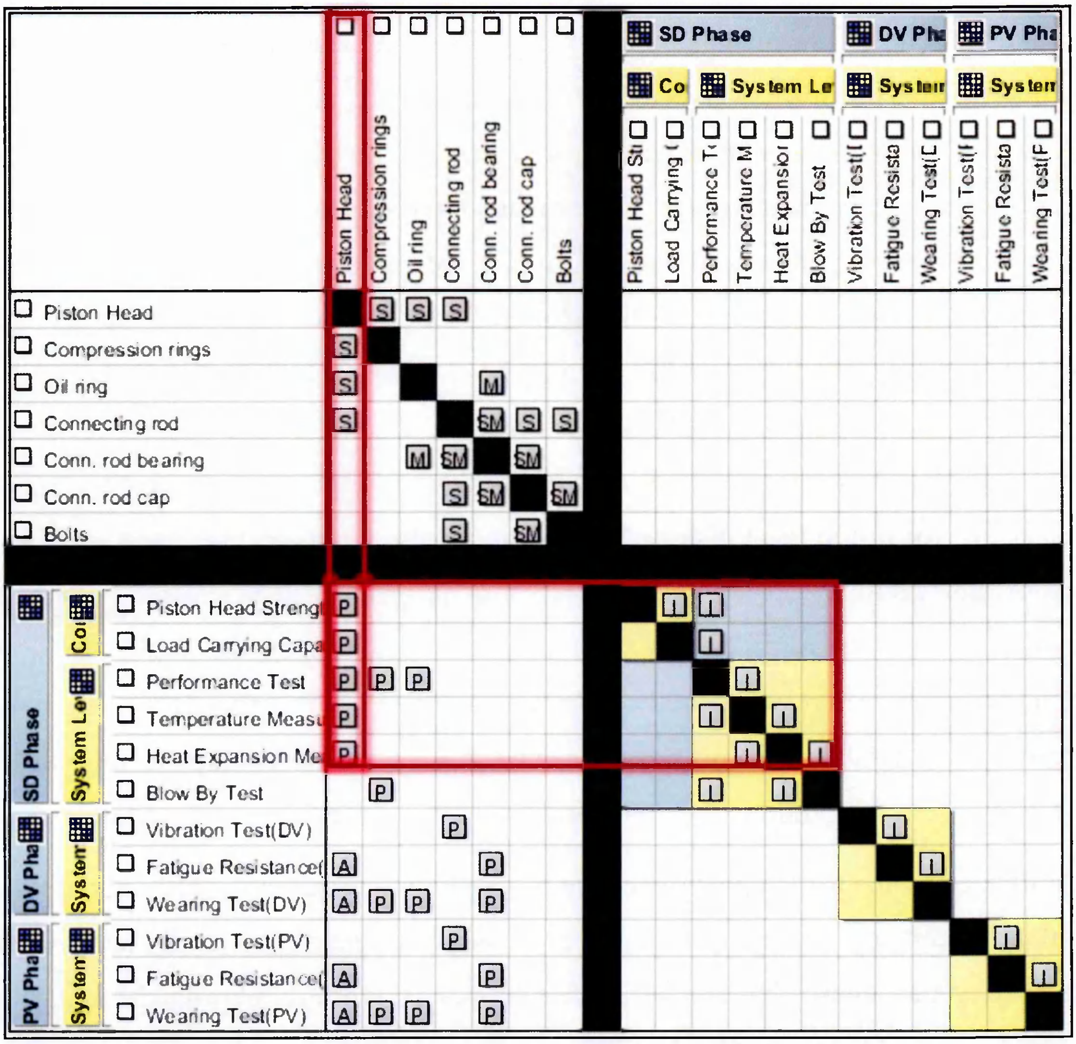


Figure 7.10 Analysing multiple domains in MDM

Jarratt (2004) proposes a DSM based modelling approach to predict the product change propagation and its effects. His research is principally focused on changes of components. He has used CDSM to capture the linkage between components of engine. If there is a linkage between two components, the likelihood of propagating a change of a component to another was analysed together with the impact of the propagated. The algorithm developed by Keller et al. (2009) was used to identify the components most susceptible to and most likely to influence change. Similar method can be used for analysing ADSM and DMM.

This research did not analyse the impact of change propagation, rather the MDM matrix was used to understand the flow of information between the domains of analysis. This MDM matrix itself provides the understanding of the company's current practice especially the visual expression of the logical arrangement of these domains. Further, and significantly it enables the engineers to reflect on their own practice.

The DMM provides information about which tests are performed on which component and clearly marks the primary and auxiliary tests to help engineers to focus on these tests. For example, looking down the column 'piston head' of the DMM (see Figure 7.10), one can easily see that it goes through five primary tests and four auxiliary tests. Among these primary tests, "piston head strength" and "load carrying capacity" are component level and others are system level tests. Scanning through the ADSM, the sequence of these tests can be easily realised. By analysing both ADSM and DMM matrices, one can realise the effect of failure in a test. For instance, both piston head and connecting rod are tested in the "fatigue resistance" test, but it is a primary test for the connecting rod and an auxiliary test for piston head. Failure in the primary test can mean that the whole test is cut short. For example if the connecting rod fails in this test then it is not possible to continue the test. In some cases this continuation might be possible and useful for learning about the behaviour of other components. However, this should be done in the full knowledge that this test must be repeated because the primary component has failed. Conversely a test might not be repeated because of an auxiliary component's failure.

The influence on the testing plan of change propagation through the product architecture can also be captured by comparing these DSMs. If a component or system fails to perform to its benchmark during a test, it destabilises the process. It breaks the sequential forward link. By looking at the ADSM, engineers can rethink and reorder the testing activities according to this situation. Also, for a change of a component or system, engineers need to reconsider whether previous tests of that component are still valid for the changes. It is also the waiting time to procure the new, changed, components, which will affect the organisation of testing activities.

Consequently, the ADSM can support a company's process failure modes and effects analysis (PFMEA). ADSM has the capability to show the potential process failures in terms of feedbacks to rework certain activities. This ADSM can be an important tool for creating a PFMEA.

7.4 Summary

It is important to capture the information flow between different domains of design, especially between testing and subsequent redesigning for the next phase. But this chapter concentrates on the dependency and information flow for the testing domain. The MDM matrix has been built to demonstrate the potential of using DSMs in organising the test process, rather than a complete depiction of the entire testing process.

The exercise can be summarised as:

1. Although DSM looks like a simple tool for dependency modelling but is not as simple as it appears. Deep understanding of the system is required for this modelling. There is a tendency to put too much information in a link. But to make it understandable and transferable, it is recommended to establish just a single type of information in a link.
2. The ADSM could be very useful to the company if an algorithm can be developed for calculating the propagating effect of change or failure of testing activities. This can help their PFMEA process.

In this chapter, overlapping of testing activities was not considered. The ADSM does not explicitly show overlapping of testing activities. The next chapter discusses the effects of overlapping of activities, especially design and testing activities, on test planning in the next chapter.

Chapter 8 A method of improving overlapping of testing and design

This study identifies that, due to long procurement times and lengthy physical tests, companies may have no choice but to start subsequent redesign tasks before testing results are available in order to meet product delivery deadlines. This increases uncertainties in design.

There is a large amount of research on optimal overlapping policies and techniques (as mentioned briefly in the literature review in Chapter 2). However, there is a little academic research on overlapping testing and design tasks. This chapter first identifies the factors that affect the overlapping of upstream testing and downstream design tasks in the company, and then proposes a method of integrated virtual and physical testing to support the overlapping the testing and subsequent redesign phases of product development.

8.1 Key concepts of overlapping activities in product development

A study by Krishnan et al. (1997) is particularly relevant in setting the context and the background for overlapping design and testing tasks. Krishnan et al. (1997) have developed a model, which categorises overlapping based on two key concepts: 'upstream evolution' and 'downstream sensitivity'.

8.1.1 Evolution

Evolution is measured for upstream tasks in an overlapping process. According to Krishnan et al. (1997), the exact value of a parameter X at time $t_i < t_f$, where t_f is the finish time of the task, is unknown, at the beginning of an activity, but the initial interval $\{a_{in}, b_{in}\}$ of the parameter X, such that, $a_{in} < X < b_{in}$, is known to the engineer (see Figure 8.1). As the activity progresses, engineers aim to reveal more information about the actual value of the parameter. The term evolution refers to the refinement of the upstream generated information from its preliminary form to a final value, which is a point within the initial interval (Krishnan et al. 1997).

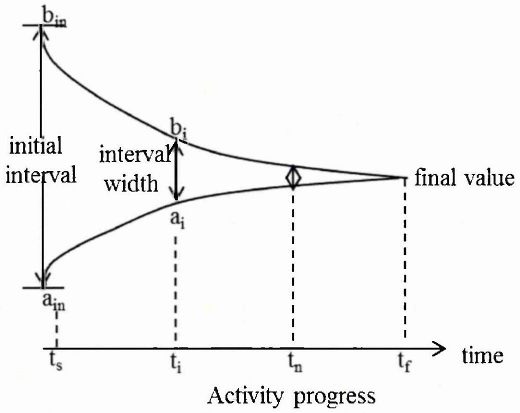


Figure 8.1 information evolution of an activity to find the final value of a parameter (adopted from (Krishnan et al. 1997))

The term 'degree of evolution' is described as the rate at which information is progressed from the start of an activity through the completion of the activity. There are two extreme cases; 'fast evolution' and 'slow evolution'. When the process fast approaches towards the final values and can achieved early in the process- this is called 'fast evolution', whilst 'slow evolution' occurs if most information evolves slowly towards the end (see Figure 8.2). Details can be found in (Krishnan et al. 1997).

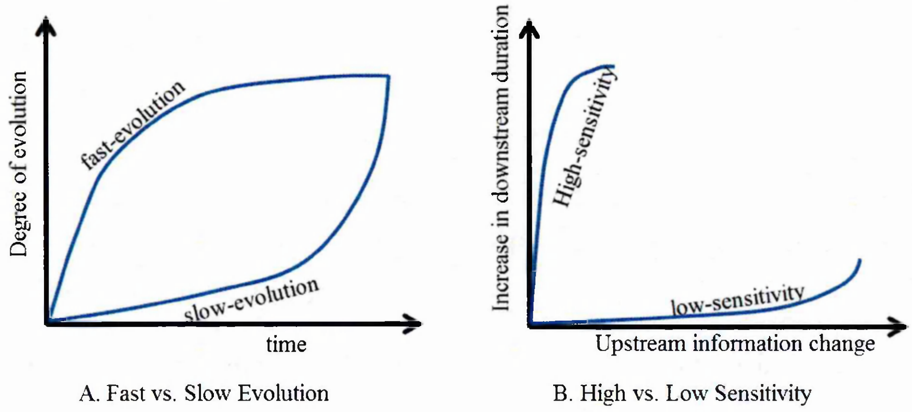


Figure 8.2 Extreme values of Evolution and Sensitivity (from (Krishnan et al. 1997))

8.1.2 Sensitivity

The sensitivity is measured for downstream tasks (see Figure 8.2B). A downstream task is sensitive to changes in the output of its upstream tasks. In 'low downstream sensitivity', substantial changes in the upstream tasks can be accommodated readily (in a short period of time) in the downstream activities. 'High downstream sensitivity' happens when small upstream changes require large amounts of rework in the downstream activity.

Krishnan et al. (1997) propose a conceptual framework, which shows four extreme situations that can arise when the upstream evolution is fast or slow combined with when the downstream sensitivity is low or high. Their analysis concludes that the case of 'fast evolution' and 'low sensitive' is the most favourable for overlapping and 'slow evolution' and 'high sensitivity' is less favourable. This classification of the exchanged information on the two dimensions, evolution and sensitivity, is useful for analysing overlapping decisions.

8.1.3 What is revealed in testing activity?

Krishnan's assumption of the term 'evolution', which refers to the refinement of the upstream information from its preliminary form to a final value, which is a point within the 'initial interval' (see Figure 8.1), is applicable for design activities but this statement can be too simple for testing activities. This is because testing activities do not refine but reveal the value of a parameter.

Engineers minimise the initial interval through design and analysis and estimate the possible value of a parameter. Hence, the expected value of a parameter is known to the engineers before commencing a test. A test reveals the actual value of that parameter. For example, design engineers assumed that an engine will produce power between 190-195kW at 2200 rpm and Figure 8.3 represents the expected phenomenon of power vs speed of an engine. A test identified that the actual power the engine produces is 192 kW at 2200 rpm, which is within their expected initial interval (the term 'initial interval' is same as Krishnan et al. (1997), see Figure 8.1). But a test can reveal that the power it actually generates is 179 kW at 2200rpm, which is outside of that initial interval.

Design processes tend to converge the information of a parameter from its preliminary value to an estimated final value. The intention in testing is to reveal the exact value of the parameter, which could be within the estimated values, if the design is successful, or could be outside the estimated value, when design has flaws. Usually, engineers will allow some margin in these expected values, for example, in the example above they defined a variation of ±2 kW as acceptable.

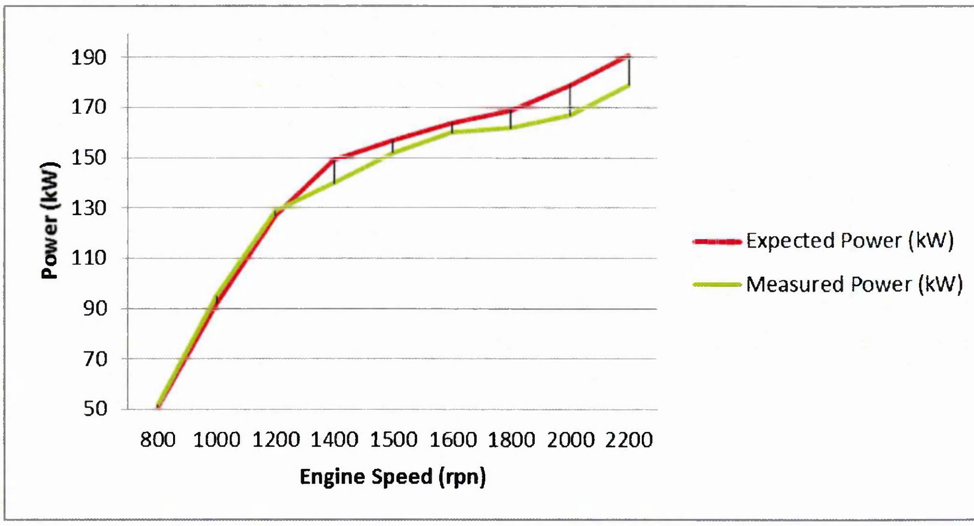


Figure 8.3 Engine - expected power curve and measured power curve

If design and testing are bundled together and considered as an upstream task then Krishnan's assumption is applicable. But if testing is considered as upstream task and design is the downstream task then overlapping of these tasks needs careful analysis. Because, the expected information from upstream testing tasks is different to that from the expected information from design tasks. In testing, engineers want to corroborate design, generally, is a case of knowing if a design can be accepted or not. If design is not accepted, more importantly, engineers need to identify how much a measured value deviates from an expected value, so that it can be improved in downstream design iterations.

8.1.3.1 Deviation from target

'Deviation' is the difference between the expected and an actual measurement, at the time of assessment (see Figure 8.4). During a test, if the engine produces any value under or above the expected values (including margins) then these results are not acceptable (see Figure 8.4). Similarly, for a "Deterioration Factor" test, engineers know that the performance of an engine will change over the time and they have an acceptable margin for each time. For example, engineers will know how much they expect the product to deteriorate after 200 hours or 500 hours of running the test. If the product deteriorates below an allowable limit, or margin, at that time, then it is not acceptable and deemed under-designed. If an engine performs above the margin then it is assumed to be over-designed.

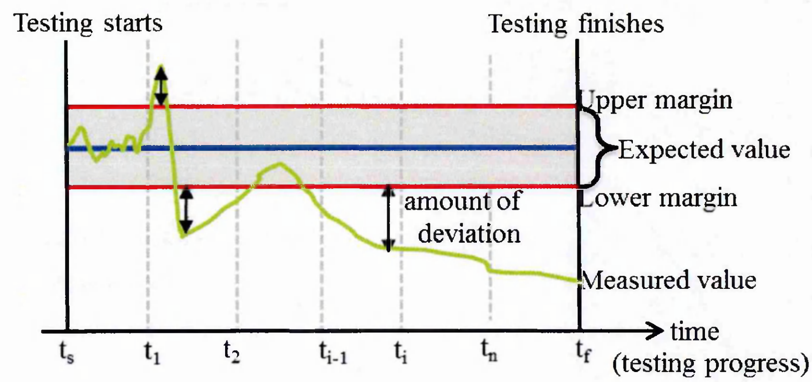


Figure 8.4 A schematic of expected value and measure value and the deviations in these values

With current levels of supporting design technologies, engineers can estimate the behaviour of a product accurately and they know what they are expecting from a test. But through a test, they want to justify their estimation. Figure 8.5 shows a schematic, where the expected values are drawn in a red line and the measured values, during a test, are drawn in a green line. For simplicity, margins are ignored. This figure depicts a case of under-design, and the product performance is gradually degrading. Also, this model assumes that the deviation is monotonically increasing (see Figure 8.5).

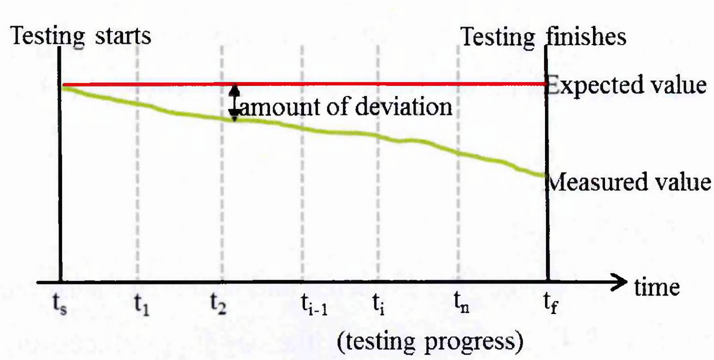


Figure 8.5 A simple model to represent deviations, over the duration of the test, from the expected value of a physical testing

In this model, it is presented that a physical test starts at t_s and finishes at t_f. Since design can be accomplished using the best knowledge available at that stage, at t_s, it is assumed that the design meets the target (red and green line meet at t_s). During the testing process, test measurements are taken and actual value of a parameter at any point is identified. Any deviation, at a point, can be identified by analysing the difference found in the test measurements from the expected value. For example, the intermediate test measurement, at t_2, identifies a deviation from the expected measurements (i.e. deviation is the difference between red and green line at t_2). During the testing, engineers will learn these deviations as the test measurements at t_1, t_2 and so on are taken. The difference between two test measurements at different times - can reveal the 'degree of evolution' (as known from (Krishnan et al. 1997)). When a test is finished, at t_f, the final measurements will reveal the final value of the measured parameter, therefore, the total amount of deviations will be known.

During the testing process, a test measurement, at any point, can therefore reveal two elements: a) actual value of a parameter, therefore how much it deviates from the expected value can be known, b) the degree of evolution, i.e. how fast the result is approaching towards the final value.

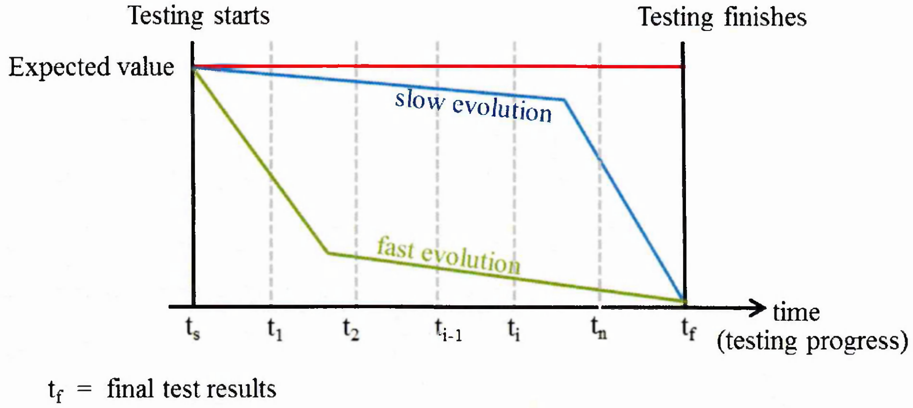


Figure 8.6 Two extreme cases of slow vs fast information evolution in physical testing process

Figure 8.6 shows two extreme cases of information evolution in testing. If the difference between two measurements is large at the beginning, at t_1 and t_2, and very small at t_i and t_n, then the testing process is 'fast evolving' (green line in Figure 8.6). In a 'slow evolution' test, there will be small difference between the measurements at t_1 and t_2 and large at the end at t_i and t_n (blue line in Figure 8.6). At the end of the process, at t_f, the final value of the parameter and actual deviation can be known.

Ideally, after finishing the upstream testing, downstream design tasks should start at t_f. However, often, the company is forced to start downstream design earlier, at t_n, for instance, while the upstream testing is still running. This may not a significant problem in a case of a fast evolution testing (green line in Figure 8.6), because difference in testing measurement between t_n and t_f is not significant, therefore, this may not largely affect the downstream design tasks. However, in a slow evolution testing (see blue line in Figure 8.6), testing measurements significantly change after t_n, hence, design starts at this point with huge uncertainty.

8.1.4 Overlapping upstream testing and downstream design tasks

Figure 8.7 presents the testing process (plotted in Figure 8.6) outlined on the gateway stages. Here, a test starts at System Demonstration (SD) phase but is completed at Design Verification (DV) phase. The gateway review takes place before finishing the testing tasks at t_n. After this gateway review, at t_n, engineers must start the design of the DV phase to meet the schedules of the next phase (see Figure 8.7). At t_n, if test measurements (i.e. the red round circle, at t_n, in Figure 8.7) are taken, engineers know the deviations. The downstream design activities try to minimise any deviations and improve the design to achieve the expected value in next iteration.

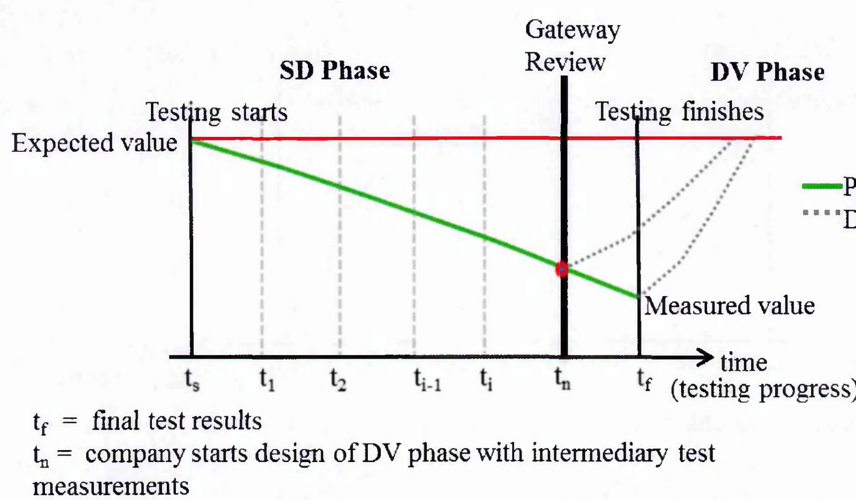


Figure 8.7 The red circle shows the starting point of downstream design tasks before finishing upstream testing

The dotted lines, in Figure 8.7, present the design works as gradually improving towards the expected design. But at t_n, engineers need to make assumptions of the final value of a parameter and the deviations that might occur at the end of the test at t_f. Usually, these assumptions are based on engineering judgement, often just looking at the progress of the intermediary tests results. Any assumption of the final value is largely uncertain if a significant change in the value of the parameter (i.e. slow evolution) happens after t_n. As mentioned before, any uncertainties in these assumptions might require considerable rework in design.

A major challenge to the company is to reduce the length of a physical test so that physical testing can be finished in gateway stage schedules, i.e. at t_n, or if not, to predict the final value of a test at this point with reduced uncertainty. The implication is that the downstream design should not suffer significant rework.

This research has proposed a method such that the uncertainty in predicting the final value of the parameter can be reduced and can support the downstream design tasks.

8.2 A method of overlapping testing and design activities

To reduce the uncertainty in overlapping, there is a need for a mechanism that can accurately predict the final value of a parameter faster than the physical testing. This research identifies that "virtual testing" can act as that mechanism.

In this proposed method, virtual testing is planned to be performed in parallel with physical testing, so that virtual testing can take the intermediary test results and process them in a form that is practical to start the downstream design tasks. The method of overlapping the upstream testing and downstream design, with aid of virtual testing, is described in the following sections.

8.2.1 The method

A process of predicting the final value of a parameter, at an earlier point, through parallel physical and virtual testing, happens in two steps:

1) virtual model is calibrated and validated through physical test measurements and
2) the prediction of final test results through simulation of that virtual model.

8.2.1.1 Step 1: Virtual model is calibrated and validated through physical test measurements

The virtual testing is described in Section 5.1.3.2. Initially, the measurements of a virtual testing can vary from a measurement of a physical testing for several reasons: a) the virtual model is not created accurately, b) theories or assumptions are not right and c) the model is not calibrated or validated due to lack of practical data. Therefore, a physical test can only be assisted with virtual testing if a virtual test is created and validated accurately, hence, will be able to predict the physical test measurements accurately.

The proposed method of virtual model validation depends on several conditions:

(i) the supportive virtual model is modelled accurately,

(ii) the model is calibrated and validated accurately with practical test measurements, and

(iii) necessary and sufficient test measurements are gathered to have a confidence in test measurements.

To create a virtual model for test, engineering experience, prior understanding of the product, previous product testing and historical data - all contribute to the model for virtual test. The case study company's virtual models are founded on the years of expertise of engineers and the capability of the software development team. The teams have the expertise to formulate mathematical models for the interacting engine components, write appropriate numerical solution algorithms, and integrate the resultant programs into workable analysis. However, new requirements and use conditions raise questions about the extent to which the model has been created accurately with the previous product information. In such a case, virtual models require modifying against the values gained from the current physical tests. Therefore, this research proposes that the variations in simulated and physical testing results can be minimized by accurate virtual modelling whilst validating the model with practical physical test data. The process of virtual model calibration and validation are discussed below.

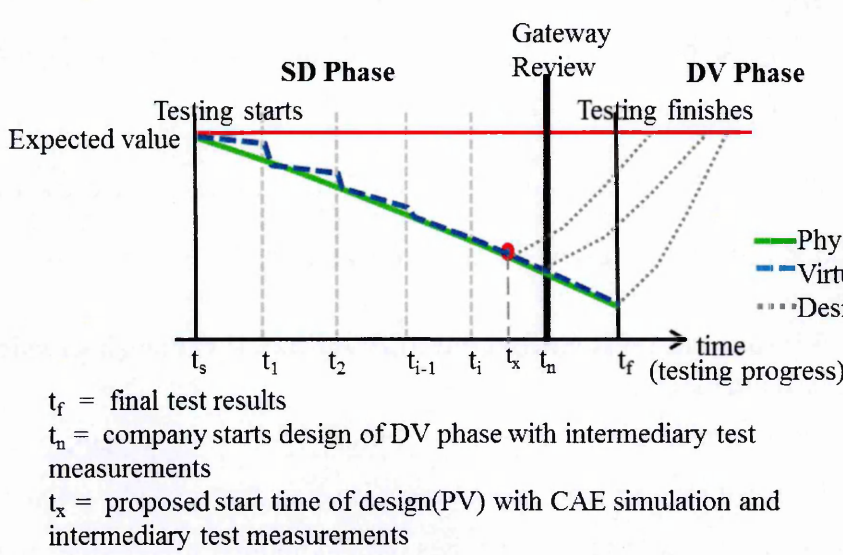


Figure 8.7 Parallel executions of virtual and physical testing to start downstream design tasks

In this proposed method, the simulation of virtual model should start in parallel with physical testing at t_s (see Figure 8.7). For a test running for a specific duration, the company takes measurements from physical tests at several set points, for example, at t_1, $t_2 \ldots t_n$ and so on. The simulated results of virtual testing should also be collected at equivalent points. That is, if the test measurements are taken after 150 cycles, for example, which took twenty four hours of running the physical test, the simulation results also have to be collected after equal amount of cycle (i.e. 150 cycles), which might take only two hours. Consider, at t_1, the physical test provides the first measurements of the parameter, which would be the practical values of the current product under test. These measurements then will be available to compare with the simulated results, considering that both were running for same number of cycles. These initial test measurements will indicate the product's behaviours and consequently the type of analysis that will be required, for example, linear or nonlinear. Therefore, the virtual model will be adjusted and improved according to these measurements.

Further simulation of the virtual model produces the values according to these measurements and can be compared again with the next test measurements at t_2. Any variations in simulated results will require to be adjusted according to test results. This process could directly help calibrate and verify the simulation results of the virtual model. In a number of iterations, the virtual model will be adjusted and improved until the simulated results are representative of the physical test results. If it is assumed that, at t_i, simulation predicts the testing measurements accurately then at this point, the virtual model is calibrated and validated with the current test measurements.

However, engineers also need to take a decision about whether the virtual model is validated and calibrated with the necessary and sufficient test measurements. At a point, say t_{i-1}, the virtual model can produce the physical test measurements accurately, yet the engineers might find that the test measurement data is not sufficient enough to validate the virtual model. So, they require a wait until a point where they feel sufficient and necessary data are available to calibrate and validate the virtual model. This means that the test needs to be running for certain amount of time to produce useful results, in which the engineers have confidence that they can predict the behaviour and value of a parameter. For example, engineers can gather enough data and can be confident on test measurements at t_i (as shown in Figure 8.7). Therefore, up to that point, the model is required to calibrate with the test measurements. At this point, the virtual model is validated and calibrated with sufficient and necessary practical data and ready to use for the Step 2 of the process.

8.2.1.2 Step 2: The prediction of final test results through simulation to start downstream design

Through the process of calibrating and validating with actual testing measurement (as described in Step 1), the virtual model should be mature enough to accurately predict a product's behaviour. The results from the virtual model simulation combined with engineering judgements of these simulated results would contribute to the accurate prediction of the final test results.

To start a downstream design task at t_n, an accurate prediction of the final values (i.e. the value at t_f) of the measured parameter are required to minimise the significant rework in downstream tasks. For example, at t_n, there are still 1000 cycles of physical testing to run, which will take certain amount of time. But the same number of cycles can be simulated in virtual testing faster than the physical tests. After the point t_i, when the virtual model is validated, virtual testing should be able to simulate the remaining test cycles (i.e. remaining 1000 cycles) significantly faster than the physical test.

Often engineers cannot predict the final value of a parameter because tests are often aggregated. Usually, virtual models are focused on an individual tests or individual parameters and are more controlled. The information from virtual tests can be disaggregated into cycles of a specific test, for example. Parameters can be analysed individually if required and can be transferred to the downstream tasks. This can also support the engineers' decisions.

Therefore, at any point after t_i, the virtual model could simulate the rest of the physical test and could predict the values of the parameter at time t_f. In this way, the uncertainty about the prediction of a final value of parameter, at an earlier point, will be reduced. If the uncertainty in predicting the results of upstream testing can be reduced then the downstream design tasks can be performed more accurately. Therefore, rework in downstream design will be reduced accordingly (sensitivity will be reduced).

In this process, an engineer might decide to start downstream design even earlier, at t_x (see Figure 8.7), when the virtual model is accurately validated with sufficient and necessary test measurement data. At this point, the company will be able to start subsequent design tasks for next phase with a simulated test result which will provide better accuracy than just assuming the values. Therefore the rework time in design is likely to be reduced.

Now the question arises, if most of the changes happen after t_x (i.e. in a slow evolution), will this virtual model be able to simulate that results. The answer is, yes, for the case study company. As mentioned before, this company has a long history of developing engines and testing the engines. They understand the product and their testing procedure, because, most of the test procedures have been run for many years (as confirmed by Engineer 1). They usually recognise the slow evolution tests and the point when the most of the changes happen in a test (this will be clear in the example illustrated in Section

8.3). With the help of virtual testing the engineers will be able to decide if they need to wait until that point to be reached or they could gain enough information through simulation and can decide at an earlier point.

8.2.2 Implication for overall product development duration

This section discusses the implication on duration of introducing the proposed method. First, general notations and conditions of overlapping are described to set up the background and the next section the implication discussed.

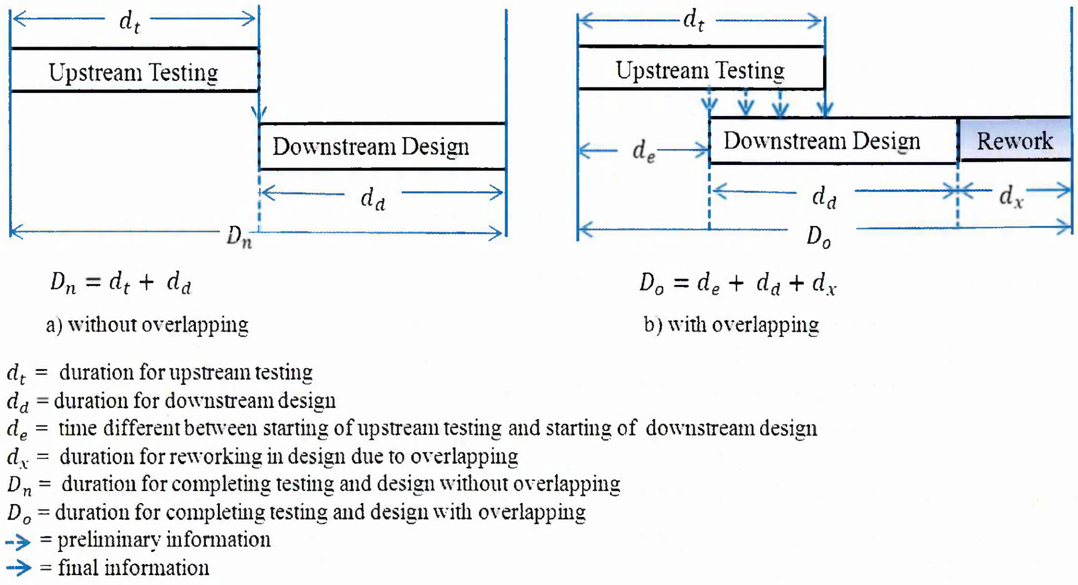


Figure 8.8 Concepts of overlapping and notations

In Figure 8.8, it is considered that upstream testing and downstream design durations are d_t and d_d respectively. The total duration of these tasks is $D_n = d_t + d_d$, when overlapping is not applied (Figure 8.8a). In general, in an overlapping situation, the downstream task can start any time after the upstream task starts and before finishing the upstream task, thus $d_e < d_t$ and $d_e \neq 0$, where d_e is the elapsed time between starting time of upstream testing and starting time of downstream design (see in Figure 8.8b). As downstream tasks start with preliminary assumptions from the upstream task, some of the downstream tasks might require reworking, when upstream task changes. The duration of rework is d_x. When the overlapping is applied, the duration is , $D_o = d_e + d_d + d_x$. Thus, overlapping will only be beneficial when $D_o < D_n$.

Another condition for overlapping is that the downstream task and rework cannot be completed before finishing the upstream task, as all upstream results needs to accommodate in downstream task, therefore, $D_0 > d_t$. The time saving through overlapping is: $D_n - D_o = d_t - (d_e + d_x)$. This equation shows that overlapping will provide better time saving with smaller duration of $(d_e + d_x)$ and need to maintain the condition of overlapping, i.e. $d_t > (d_e + d_x)$.

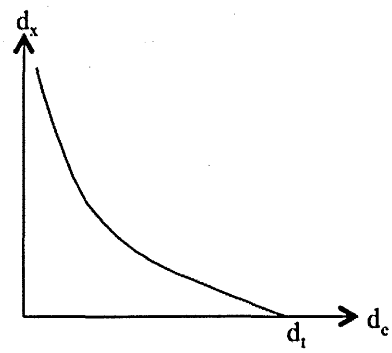


Figure 8.9 Relationship between d_e and d_x

Delaying the start of the downstream task, i.e. increasing the d_e, will allow accumulating more upstream results in a downstream task, which might mean less time is required for reworking. Therefore, d_x tends to decrease with the increase of d_e (as shown in Figure 8.9), and $d_x = 0$ at d_t, when downstream task starts after finishing the upstream task (no overlapping). Figure 8.9, shows the relationship graph between d_e and d_x.

According to Krishnan et al. (1997), this relation of d_e and d_x can be affected significantly, when a small change in the value of the parameter of upstream task causes significant downstream work - thus increasing duration d_x and see section 8.1.4 above for more details. Therefore, a critical decision needs to be made about when to start the downstream tasks, i.e. d_e, so that downstream design doesn't need significant reworks (i.e. minimum d_x).

This research suggests that downstream tasks should be started once the virtual model is accurately validated with sufficient and necessary testing measurements after t_x (see in Figure 8.7). So, d_e is the duration between the starting time of the test and the time required for virtual model validation (i.e. duration between t_s and $t_{x,}$ in Figure 8.7). After this point, the downstream design tasks can start with reduced uncertainty that means minimum rework will be required.

With recent improvements in CAD and CAE tools, downstream design changes/rework and analysis can happen in a shorter time. Also, virtual models of one phase can be used to optimise the design for subsequent phase. Therefore, the sensitivity of the downstream design, that is d_x (duration of rework), can be significantly minimised through CAE analysis. The downstream design sensitivity can also be minimised through the effective communication between test engineers and design engineers. Other factors such as the products' modularity, robust design, and anticipation by downstream designers for changes in upstream information, can all help reduce the sensitivity of downstream design (Krishnan et al. 1997).

8.3 Example 1: Gross Thermal Cycling

This section illustrates how the relation between d_e (staring time of downstream design) and d_x (rework time due to overlapping) might change, if the proposed method is applied, through an example.

A test for gross thermal cycling is an example of a lengthy endurance test, which checks the fatigue resistance of the cylinder head. This example was chosen because, frequently this test runs over the gateway stages and engineers need to start downstream redesigning while this test is still running. This is a critical test because it is performed on a core component; any changes of this component hugely impact the total engine system. Also, this test is very costly to run. This test is usually planned for DV phase, at least three times for three variations of engines, but often needs to repeat at PV phase to validate any remaining hardware variations. In recent projects in the case study company, it has become a norm to repeat this test at PV phase.

First the background of the test is discussed and then it is demonstrated how the behaviour of d_e and d_x could change when supported by virtual testing.

8.3.1 Purpose of the gross thermal test

This gross thermal test is a procedure for determining the thermal fatigue resistance (see Appendix A for more about fatigue) of core engine components, by subjecting the engine to controlled, rapid coolant temperature change cycle. The cycle is normally applied to evaluate the cylinder head and cylinder head gasket. However, other engine components are subjected to this gross thermal test. Each test cycle is structured to run for 7 minutes (420 seconds) cycle and at least 8500 cycles must be achieved. This equates to approximately 1000 hours of test and means that the engine is in test bed for at least two months (8 weeks). The objective of this test is that when an engine is run for extended periods (1000 hours) in the test cycle given in this specific procedure, it will simulate the conditions that the engine will meet in service over the full lifecycle. The testing team records the data stream from the physical set up at several set points as the physical testing progresses. Cycle adherence is checked and sensors reading are taken every 24 hours. Test measurements are recorded every day for this test. But, the engine is finally checked once the test is finished. The actual examination of the engine will range from simple visual inspections to accurate measurements of degradation of a given characteristic i.e. wear of a component surface or rate of change of performance, leaks and cracks.

8.3.2 Changes in durations d_e and d_x with parallel virtual testing

This example was created with the engineers (Engineer 1 and Engineer 3) in the case study company. This example was the result of several discussions. The engineers understand that there are significant amounts of overlapping in their process. They also realise that overlapping the downstream design while upstream testing is still running is the most critical one. But they don't have a method of managing this overlapping. As mentioned before, redesign for the next phase of product development must take place after a gateway review. They will acquire as much information as they can from the upstream testing by observing the pattern of test measurements as well as through engineering judgements. They will have meetings with test engineers, product engineers and senior validation managers to decide about the test and will release the information to the downstream design team.

It seems, there is little that is acknowledged by the engineers about the academic research in overlapping. None of the overlapping method that was suggested in literature is known to be applied in the company. The terms 'evolution' and 'sensitivity' were not used by the engineers, during the discussions. However, they use the term "deviation" when they explained the testing process.

During the discussions the author did not use these formal terms to describe the problem but simpler terms and drawings were used. The inquiry questions were rather simple, for instance, "when will you start the redesign for DV phase" or "how long it will take to rework if there is a deviation".

Whenever engineers start a test, they clearly know that they have to start the downstream design tasks after the gateway review. That is d_e can be estimated. But the rework time is often described as "long" or "can be relatively short" or "can change very quickly". Any qualitative measurement of d_x is not a case in this company.

For the purpose of producing this example, engineers were asked how the behaviour of d_e and d_x will be observed in a regular case. This is shown in Figure 8.10(curve a). The horizontal axis represents the elapsed time, d_e, i.e. the time between starting time of upstream testing and starting time of downstream design. The vertical axis represents the time that might be required for rework in downstream design. The duration is presented in days, which was easy for the engineers to estimate. Engineers identified previously that the test does not produce any significant results during the first 4 weeks and most of the fatigue of the components starts to appear towards the last two weeks of the test. It is not recommended to start the redesigning the parameter of interest, before 4^{th} week, thus d_e is at least 28 days. That is why the horizontal axis starts at 28days and up-to 56 days, i.e. the total duration of the test (8 weeks). If they start redesigning after 28days, they might need as long as two weeks to accommodate any change identified in the test after that period. During the last week (i.e. after 7000-7500 cycles), they can decide more accurately about the product's behaviour. However, they can only do final investigation after completing the test and many unexpected phenomena might appear which was not possible to predict before. This might cause a significant rework, can take as long as d_d, i.e. total duration for redesign, resulting in large increases in total design time (i.e. D_0 becomes larger then D_n, see Figure 8.8).

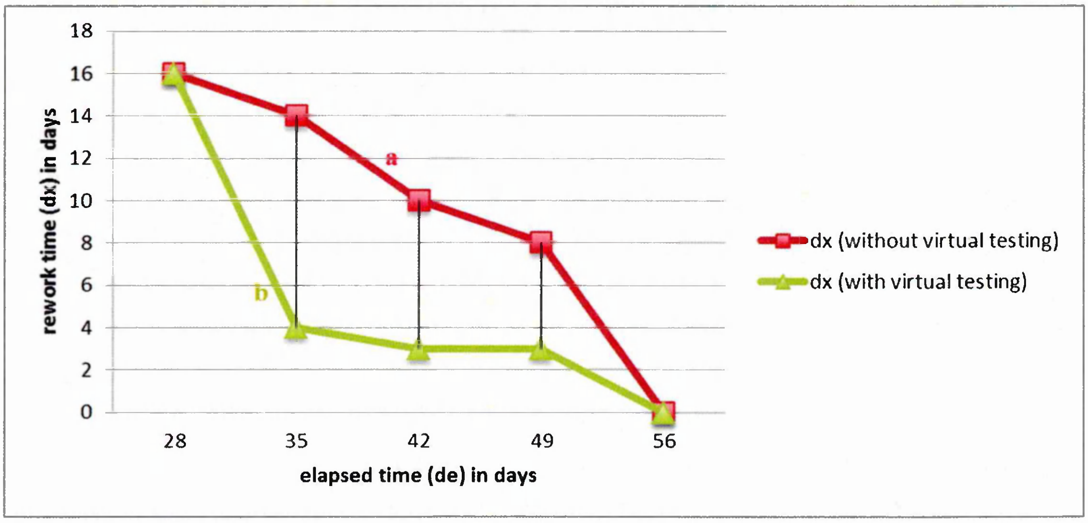


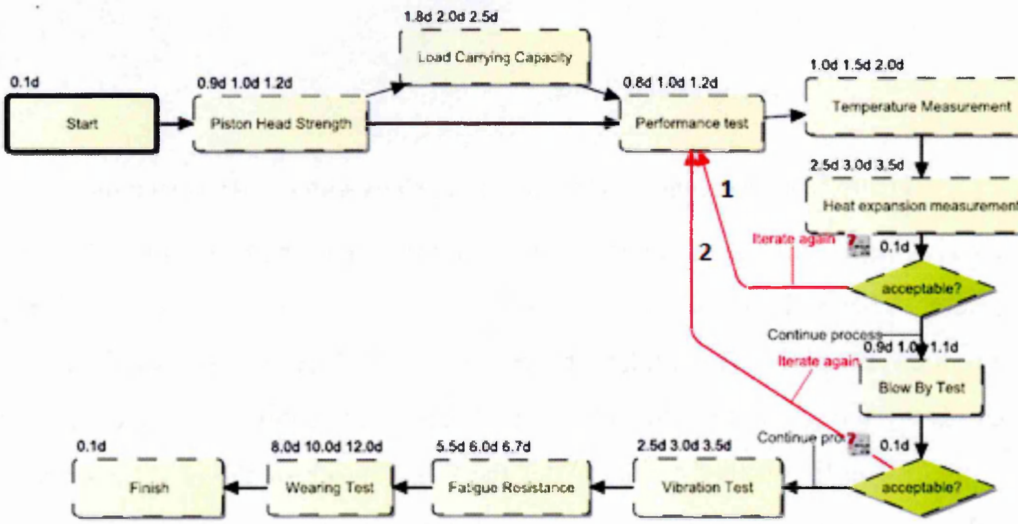
Figure 8.10 the change in behaviour of de and dx with use of virtual testing

The potential of using virtual testing and how this might affect the behaviour of d_e and d_x was considered. This is shown in Figure 8.10(curve b). After four weeks, the physical test starts to produce enough data to validate the virtual model, thus up to this time (28 days), the virtual testing model could be developed and calibrated, although not recommended for use in predicting the testing results. After the fourth week of the test, the engineers will be able to use test measurements, combined with historical data to virtually model the behaviour of the component under test. The virtual model will be calibrated and validated with the physical testing measurement in several iterations using the test measurements, which are taken daily in this test. Thus it can be assumed that the virtual model will be capable of simulating the test results after 5th week. To run a simulation will take a day at maximum as only 2000-2500 cycles will be remained to run after 5th weeks. Therefore, the subsequent design can be started any time after 5th week.

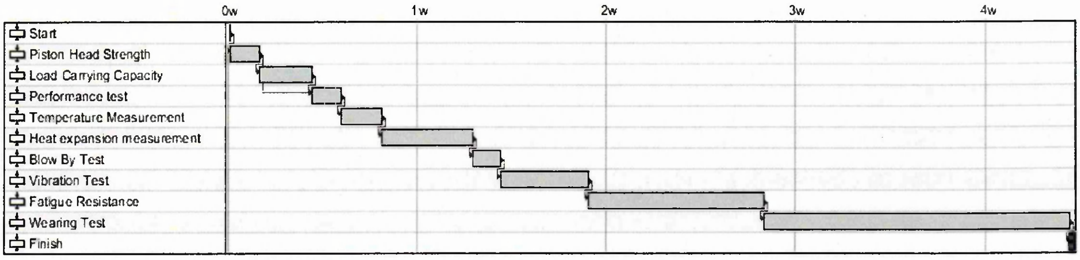
At this point, the engineer might decide to wait for gateway review or possibly start the downstream design tasks even earlier. By providing accurate data, the design might not suffer substantial rework. As mentioned before, virtual testing of one phase also assists the CAE analysis of next phase. As design is assisted by CAE analysis, any changes in design can be done in considerably shorter time. Therefore, the duration in downstream design rework, d_x, can be reduced substantially with the proposed addition of virtual testing. Learning from the parallel virtual testing will also reduce the uncertainties in procurement.

8.4 Example 2: Group of tests

In the previous example, overlapping between upstream testing of a stage and the downstream design of a subsequent stage has been considered. Often testing activities are overlapped in a single stage in order to finish tests within the gateway stages. The aim of this example was to explore the effect of using proposed method of parallel virtual and physical testing on the overall duration of a group of testing activities. In this example, a group of testing activities is considered that has been modelled in ADSM in the Chapter 7 (Only the SD and DV level tests are considered). Figure 8.11 (a) shows a flow diagram of these tests.



(a)



(b)

Figure 8.11 (a) Flow diagram and (b) Gantt chart for ADSM presented in Figure 7.9

In an ideal case, these tests should be planned to be performed sequentially as shown in Figure 8.11 (a). However, often the company cannot perform all these tests sequentially because some of the mechanical tests like fatigue resistance and wear tests take a significant amount of time, can be seen from the Gantt chart in Figure 8.11 (b). In such a case, the company might start these lengthy tests earlier with some assumptions on previous tests. For example, in this example (see Figure 8.11 (a)), the fatigue resistance test requires input from a vibration test. To allow the fatigue resistance test to start earlier and thereby allowing it to fall earlier in the testing sequence, some assumptions about the results from its upstream vibration test are required. The same applies to the wear test. The fatigue resistance test uses assumptions about the vibration test as input, and the wear test uses assumptions about the fatigue resistance.

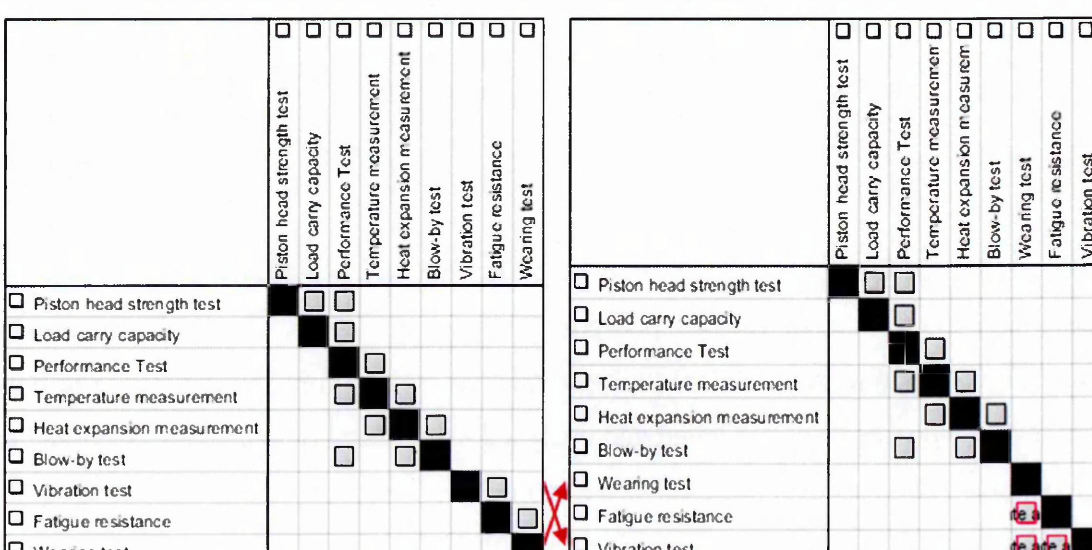


Figure 8.12 (a) The original ADSM as in Figure 7.9, (b) altered ADSM for increasing concurrency in the process

A practical way of presenting the overlapping of these tests in an ADSM is presented below. Reordering of rows in the ADSM in Figure 8.12 (a) gives the new ADSM in

Figure 8.12 (b). In the new ADSM the wear and fatigue resistance tests are considered to start earlier than the vibration test to accelerate the whole process of testing. These tests are moved earlier in the sequence; the marks below the diagonal denote making assumptions (see in Figure 8.12 (b)). Any uncertainties in these assumptions can lead to rework.

Figure 8.13 shows an example in which wastes are highlighted (in green) due to poor assumptions that were made at the beginning of wear and fatigue tests.

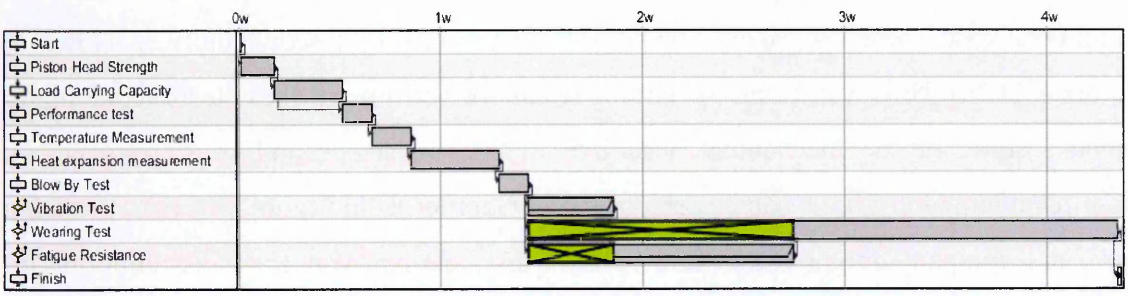


Figure 8.13 Wastes due to rework highlighted on the project Gantt chart

These tests therefore required reworks. In an overlapping process, if there are significant amount of reworks, the overall completion time can be even larger than the sequential process. This thesis argues that these effects of poor assumptions can be reduced with an appropriate use of virtual testing.

8.4.1 Modelling the revised testing activities

Currently, engineers overlap these testing tasks to reduce the overall completion time but times vary significantly between projects. Therefore, to estimate the completion time, processes are modelled and simulated.

The modelling process started with gathering the data about the ideal ("as-is") testing process and the changes to be incorporated into the future ("to-be") situations. Simulation was then used to explore the consequences of the proposed changes. The flow diagram of the process was modelled and analysed using Cambridge Advanced Modeller (CAM) tool. This section describes the modelling process briefly but a detailed description of the process can be found in (Wynn 2013).

Two scenarios were considered:

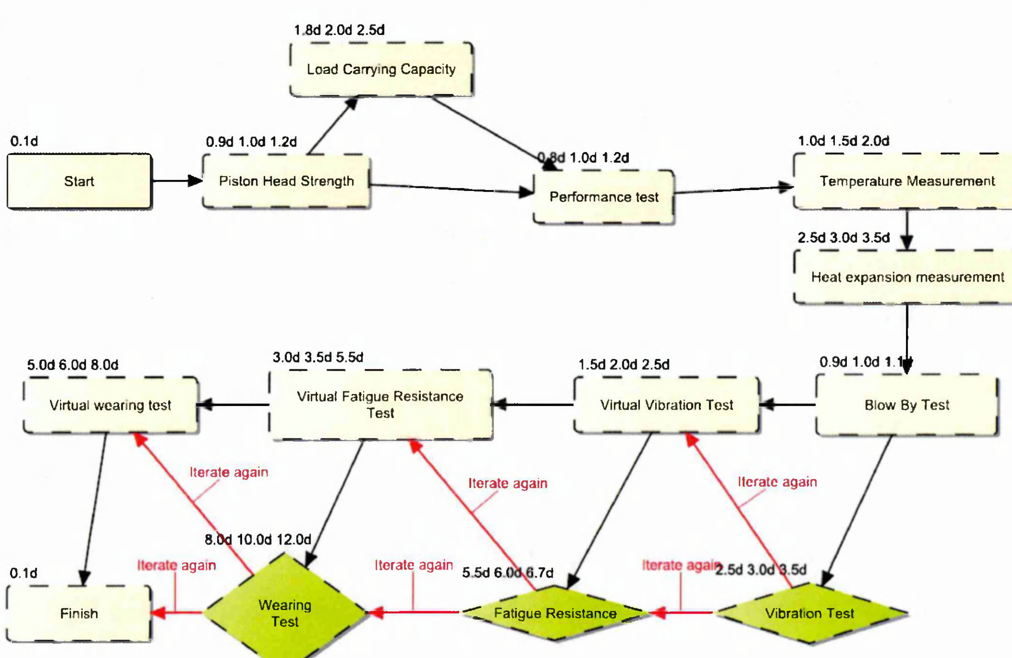
(1) ideal way of sequential testing activities (flow diagram in Figure 8.11 (a))

(2) revised way of overlapping testing activities, supported by virtual testing (Figure 8.14(a)).

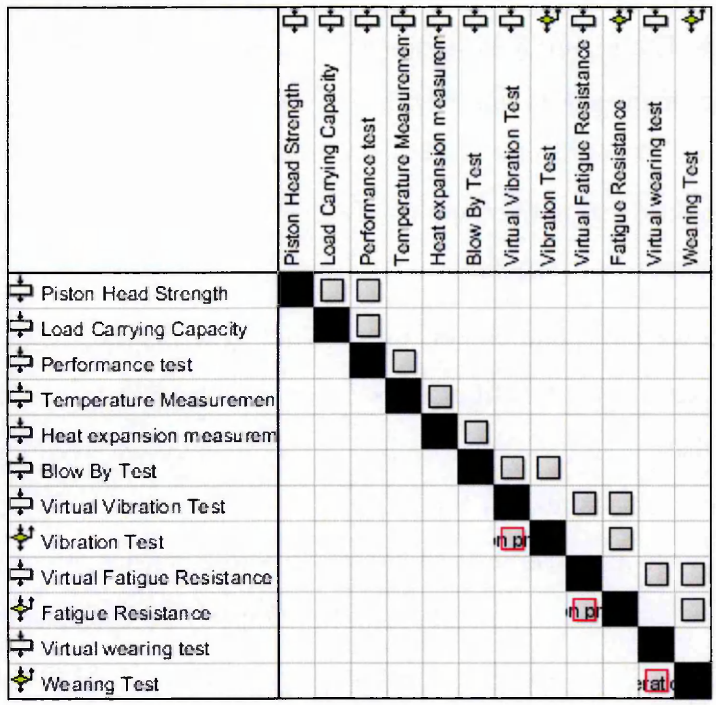
The revised flow diagram in Figure 8.14(a) depicts how virtual tests can be introduced to the testing process. The Figure 8.14(b) shows how the introduction of virtual testing can be visualised and managed in an ADSM. For simplicity, earlier iterations in the performance test, i.e. flow 1 and 2 in Figure 8.14 are ignored. This helps to focus on the iterative effects that are due to overlapping. The number of physical tests and duration of each remain the same in both scenarios.

Figure 8.14(a) shows the modelling of the "to-be" flow diagram where the vibration, fatigue and wear tests are reconfigured as iterative construct (green diamonds). That means that when these tests are finished, they 'may' or 'may not' feed the information to the successor tests. The simulation logic interprets these situations by not forcing their successor tests to wait for these tests to complete before starting. For instance, 'fatigue resistance' will not wait for 'vibration test' to complete before it starts. However, if 'vibration test' feeds information into 'fatigue resistance' later on during a simulation, then 'fatigue resistance' test will be reworked along with all its successors that had already been executed.

From the Figure 8.14 (a), it can be seen that the flow (red arrow) from 'vibration test' to 'fatigue resistance' for instance, is labelled as 'iterate again', this means that the feed will only occur in a case of error, i.e. the assumptions made by fatigue resistance to start early have turned out to be inaccurate, therefore rework is necessary. The likelihood of rework can be set in these iterative constructs.



(a)



(b)

Figure 8.14 (a) Flow diagram, (b) ADSM for testing plan supported by virtual testing

Within every activity examined in the case study, a representative minimum, expected and maximum duration was estimated for each physical test (although actual values are not presented to preserve the company's confidentiality). Their distribution was found to suggest a triangular probability density function (TriPDF). This model was also used to assign durations to the corresponding virtual tests. For instance, the duration for virtual vibration test is set as TriPDF (1.5, 2, 2.5) for the first iteration. Here, it is assumed that in a best case, the virtual vibration test can be calibrated and validated with necessary and sufficient test measurements within the halfway through of the vibration test (i.e.1.5 days). It is most likely that it will take 2 days and in a worse case, it can take as long as the total duration of vibration tests (i.e. 2.5 days). However, a vibration test will take significantly shorter time for the case of iteration. So, the duration for consecutive iteration of virtual vibration test has been set as 1 day. Also, it has been set that the virtual vibration test will not be performed once the physical vibration test is finished. Other virtual tests are configured in this logic.

As mentioned earlier in this chapter, elapsed time is the time difference between starting of an upstream activity and the starting time of a downstream activity (see in Figure 8.8). Currently, engineers decide the starting time of a downstream activity by looking at the progression of upstream tests. They also use experience and tacit knowledge. The completion time largely depend on elapsed time and how accurate the engineers are in making assumptions. But in this proposed method, elapsed time is determined by estimating the time that is required to calibrate and validate the respective virtual tests. This has been modelled by linking that the inputs from a virtual test are mandatory to start a corresponding physical test. For instance, fatigue resistance test need inputs from the virtual vibration test to start, that means, fatigue resistance can only start when virtual vibration test is calibrated and validated (see in Figure 8.14 (a)).

8.4.2 Simulation and analysis of the model

These two scenarios were then executed using 10,000 Monte Carlo simulation runs. The first used the ideal ("as-is") testing process. As the durations of these physical tests are modified, it was essential to simulate to depict sequential process duration (without overlapping). The Monte Carlo simulation of the duration of the current process yields a histogram distribution as given in the following Figure 8.15. The mean duration is 28.91 days with a standard deviation of 0.94 days for the given activities. In a best case, this process will complete in 25.82 days and a worse case it may take up to 31.98 days. The chance of completing the project on 29.79 days is 80%.

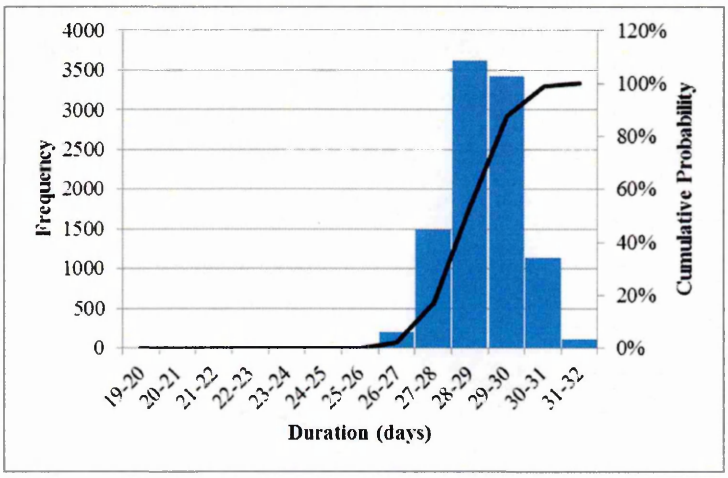


Figure 8.15 Histogram for sequential process duration

Second, the overlapping ("to-be") process is used. This experiment studied the completion time of the overlapping process by varying the probability of rework. If the engineers want to reduce the completion time to 25 days, for instance, and still want to achieve the 80% confidence of that the project will finish on time, then they will have to reduce the rework time that is due to overlapping. Typically, rework in one activity can propagate rework in other activities and higher order activities require careful consideration when the probability of rework is set. To keep this exercise simple, the propagation effects of rework on higher order activities have been ignored. Also the likelihood of rework for these three iterative construct (i.e. the vibration, fatigue and wearing tests) have been set equal. The next section discusses the output of the experiment that was performed.

8.4.2.1 Simulation outputs

The left column of Table 8.1 shows the values that were considered for the likelihood of rework of each overlapped activity. The likelihood of rework was changed to execute 10,000 Monte Carlo simulation runs in each case and is recorded in Table 8.1.

Table 8.1 Likelihood of rework

Likelihood of rework (%)	Mean (days)	Standard deviation, σ (days)	Likelihood of finishing within 25 days (%)	Likelihood of finishing within 26 days (%)
50	24.05	3.07	56	65
40	23.35	2.58	65	73
30	23.04	2.29	73	80
25	22.85	2.12	77	84
20	22.80	1.96	81	86

Figure 8.16 shows the probability mass functions (PMFs) and cumulative distribution functions (CFDs) of simulated duration outcomes when likelihood of rework is set as 30%. Figure 8.16 also shows that the likelihood of finishing within target of 25days is 73% and can be increased to 80%, if 26 days are allowed (when likelihood of rework is set as 30%).

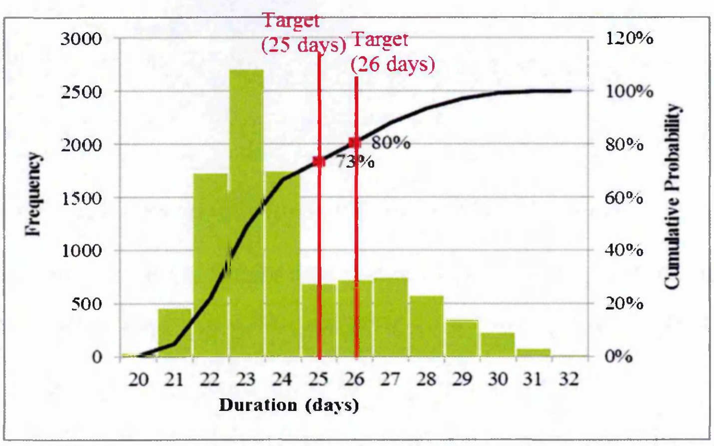


Figure 8.16 Histogram of duration PMFs and CDFs

Similarly, from the Table 8.1, it can be seen that the likelihood of rework has to be decreased to 20% to achieve the target of 25 days (likelihood is 81%) for completing the proposed overlapped activities.

Not surprisingly, this analysis reveals that if the likelihood of rework can be reduced, there is a greater benefit of overlapping. In the proposed method the likelihood of rework can be reduced by improving the capability of virtual testing. This means that if elapsed time can be increased, i.e. the time to start the downstream test can be delayed then additional time is available to calibrate and validate the virtual models with real test data, which can benefit in reducing the likelihood of rework. Eventually, this might increase the total duration slightly but there can be achieved higher confidence of completing the given activities within target time. Therefore, the real benefit of this integrated approach is the use of virtual tests to minimise the uncertainties caused by overlapping the physical tests.

Similarly Lin et al. (2008) has reported that more rework occurs as the degree of overlapping increases. Hence, engineers need to make a decision on how much confidence they want to achieve to finish a network of activities within target time, also how much delay they can allow to build up before the start of the downstream activity in a case of overlapping. This kind of simulation analysis is useful when engineers are negotiating the time and cost targets, as well as choosing an acceptable risk when planning testing activities.

Virtual model creation and virtual testing may appear to increase the costs and resources consumed in each phase of the process. In any product development, balancing cost increases against possible time savings will be of critical importance.

8.5 Modelling of cost for this method

Effective communication between physical testing team and the CAE team is a key success factor for this structure of parallel physical and virtual testing. The cost will depend on two main factors: *communication cost* and *virtual testing model establishment cost.*

As mentioned before, intermediary testing measurements are taken in set intervals, these measurements are key for developing, calibrating and validating the virtual models. These measurements are usually taken and understood by test engineers. The CAE team, who is responsible for building the virtual models, might have very little understanding of these test data. Therefore, an effective communication between test engineers and CAE engineers is required for delivering these measurements in a useful form.

Conversely, simulation results from the CAE team needs to be compared with the testing measurements. This might not make sense to test engineers. So both teams will need to communicate regularly to create an effective joint understanding. Test engineers will collect the test measurements; the CAE team will collect the simulation results. To make this effective the two teams will need to compare and decide if the virtual model replicates the actual test measurements. As discussed before, initially, virtual (simulated) and physical testing results may differ in several ways. Such discrepancies may determine the number of meetings required, which might increase with the level of uncertainty and potential dependencies (Loch & Terwiesch 1998). The number of meetings required can be used as an indication of these discrepancies and as these meetings incur significant cost can be used as a measure of cost.

Initially a fixed cost C is required to build the virtual model. But this model largely exists during the analysis phase of design. This cost will also depend on the company's capability in CAE modelling and simulation. With a well-established CAE department, this cost might be lower than outsourcing.

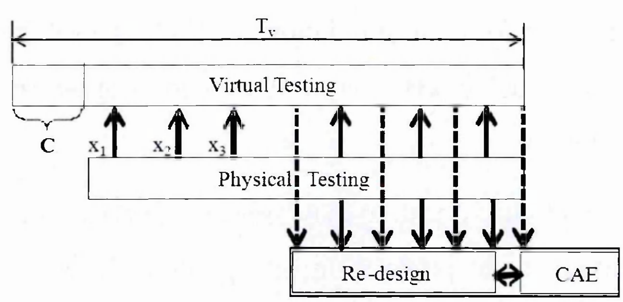


Figure 8.17 Information exchange between virtual testing, physical testing and downstream design

Assume that the cost for each meeting is X_i, for meetings i = 1, 2,.. n (as shown in Figure 8.17). After the virtual model (which is created at a cost C) is validated, the frequency of meetings can be significantly reduced. Each meeting may results in modifications and further simulation in the virtual model, at cost of Y_i. A regular maintenance and opportunity cost M is incurred per unit time, for the virtual test duration T. But, if a company has committed human resources for CAE analysis throughout the process, this maintenance might not add extra marginal costs. The cost of additional virtual testing model is measured as:

$$C_{VT} = C + \sum (X_i + Y_i) + MT \quad (1)$$

Initially, this approach of parallel physical and virtual testing will increase the cost of testing in a single gateway stage. However, the real benefit of using parallel virtual testing continues during iterations, as this might avoid extending testing into a subsequent gateway. Even with iteration, the cost of running the virtual testing phase will be approximately $\sum(X_i + Y_i) + MT$, because the model building cost C will be small as the virtual testing model is already mature. The number of meetings will also be relatively low. The duration of physical testing in this phase will be shorter, and uncertainty will be decreased in redesign. Therefore, in terms of cost and time, overall savings are likely to be improved by the proposed process.

The proposed method of parallel virtual and physical testing was validated with the senior engineer in the company. It was highlighted that this combined approach of physical and virtual testing methods had the potential to reduce iterations and thereby the number of physical prototypes saving time and cost. Therefore, the approach of virtual and physical testing integration appeared promising to the engineers of the case study company.

As this is an analytical model and any timings for virtual model implementations are only estimates, the time estimations in Table 8.1 may be unrealistic. The time required for virtual model creation depends on the company's CAE department and experience of software engineers, and depends on similar models developed. The number of iterations between virtual and physical testing will vary for several reasons such as the level of uncertainty, as well as the accuracy and completeness of communication between testing engineers, design engineer and software engineers.

Different tests benefit from integrating virtual testing with physical testing in different ways. Some benefit by focusing the tests, or identifying future values to minimise the number of iterations, while others require running for shorter periods of time. For example, in a constant speed and load situation, an engine has its quantities of fuel and air intake regulated, with the goal of achieving desired power ratings. An engine might require several iterations in design and test to achieve these desired power ratings. A virtual testing using a validated model can predict the likely consequences of certain values of fuel and air intake of the engine, thus suggesting appropriate values for next iteration in redesign and physical testing

However, not all physical tests will benefit from this approach. For example, consider a case of physical testing, where information evolves quickly and engineers can start downstream design tasks quite accurately with acceptable sensitivity. This test does not require support from parallel virtual testing. Also, there will be cases where some of the phenomenon which physical testing addresses are not effectively possible to virtually model and test, therefore this method is not applicable. Earlier in the thesis, sealing was mentioned as such a case, where although virtual models are in principle buildable they may be too complex, take excessive time or the models themselves may not be sufficiently accurate.

8.6 Implications of parallel virtual and physical testing for the product development structure

This research suggests that there will be structural changes in the company's product development process by introducing virtual testing in parallel to the physical testing in each phase of the product development process. The proposed model distinguishes virtual testing from the initial CAE analysis.

8.6.1 CAE for procurement

Using initial CAE modelling and analysis, a design team can iterate the design to develop a product that better meets cost, performance, and other constraints. CAE analyses enable the company to carry out optimisation earlier in the product development cycle as well as improving product specifications to suppliers. Clear, precise and accurate specification can reduce the procurement time (as mentioned by Engineer 3). It is often difficult to separate the design tasks and CAE tasks, because design and CAE analysis take place together in the same team. Therefore, the proposed model incorporates design and CAE analysis and suggests more iteration through CAE analysis before procurement of prototypes. Further CAE analysis will also help to set-up physical test conditions, input parameters and sensors locations for physical testing.

8.6.2 Parallel virtual testing to assist lengthy physical testing

The proposed product development process structure places the virtual testing parallel to the physical testing. Figure 8.18 represents the product development process that results from incorporating virtual testing as a separate entity. There are two aims:

(i) to improve the understanding of intermediary physical testing results which will enable the start of subsequent redesign tasks with less uncertainty,

(ii) to reduce the physical testing duration or number of iterations in physical tests.

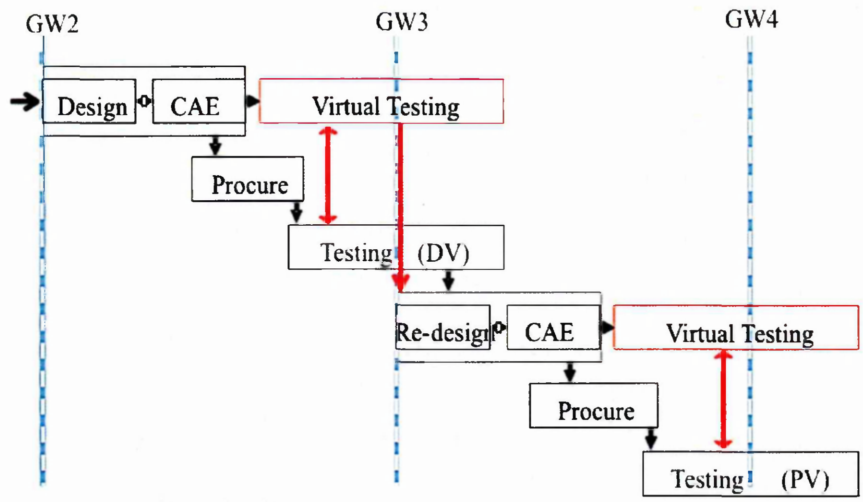


Figure 8.18 The proposed PDP structure with additional virtual testing activities

Different tests gain different advantages from integrating virtual testing with physical testing. Some benefit by focusing the tests, and identifying future values to minimise the number of iterations required to reduce uncertainty in design. Others may benefit by reducing the required running period for a physical or virtual test. For example, for constant speed and load, an engine has its intakes of fuel and air regulated, with the goal of achieving desired power ratings. An engine might require several iterations in design and testing to achieve these desired power ratings. Virtual testing using a mature model can predict the likely consequences of certain values of fuel and air intake of the engine, thus suggesting appropriate values for next iteration. When a virtual test is able to accurately predict the behaviour of the engine, then the number of physical testing hours for durability can be minimised, saving time and reducing cost.

8.7 Summary

It was identified that overlapping is one of the key issues in the case study company. It can cause uncertainties and iterations in the product development process. However, overlapping also offers opportunities for process improvement, especially in parallel virtual and physical testing, which can reduce uncertainties in necessary overlapping between design and testing activities.

Overlapping between upstream testing and downstream design happens in each stage of product development and is a characteristic of the design process.

Long-lead time for procurement and lengthy physical tests are the apparent causes of these overlapping. However, design changes are the underlying issue that affects the lead-time for procurement of prototypes.

This research suggests a model to reduce the uncertainties associated with overlapping between testing and redesign. This model introduces virtual testing in parallel to the physical testing in such a way that virtual testing can acquire intermediary physical testing results and release to subsequent design with less uncertainty, so that design suffers less rework. Also, CAE analysis has a prominent position in the model because CAE analysis can reduce the design changes and rework time.

This approach proposes changes to the structure of the product development process that reflects the roles of design, analysis, virtual and physical testing in the overall process more accurately.

Chapter 9 Conclusions and future works

This chapter concludes the thesis. The research questions are revisited and the research contributions are summarised. Then the research is validated and limitations are discussed. At the end, opportunities for future works are described.

9.1 Review of research contributions

This section reviews the research contributions under two headings: answers to the research challenges and additional contributions.

9.1.1 Response to the research questions

The success of this work in addressing the research challenges defined in Chapter 5 is considered in this section. Initial research questions were addressed in Chapter 5.

The first question concerns prioritising testing activities:

> How can prioritisation of testing activities be improved?

The case study showed that testing activities are prioritised based on risk. In FMEAs, technological risks are identified. Testing activities are used to detect the presence/absence of these risks. Testing activities which detect high risk get high priorities. But this research has identified that the current approach is not sophisticated enough for prioritisation of testing activities, which the senior engineers agreed with. Voice of customers and voice of regulations are also influencing factors on how a product should be tested. Therefore, the customer and/or regulation requirements should play an important role in prioritising the testing activities. This research suggests a method of prioritising testing activities, which integrates the outcome of QFD that evaluates the importance of technical requirements with the outcome of FMEA that evaluates the risks of these technical requirements. Engineers in the case study company evaluate this method as improving the prioritisation of the testing activities.

The second question concerns dependencies between testing activities:

> How can dependencies and relationships across components and testing be effectively captured and visualised to improve testing planning?

A conventional way of managing testing activities is in a design verification plan and report (DVP&R), which inadequately capture dependencies between testing activities and ignore the link between a component and its testing. A multiple domain matrix (MDM) based modelling approach was proposed to capture the dependencies between testing domain and component domain. Three matrices are modelled, among these, the CDSM (Component Dependency Structure Matrix) models the dependencies between components, the ADSM models the time sequences of testing activities and a DMM (Domain Mapping Matrix) relates components with associated testing activities. The proposed ADSM (Activity Dependency Structure Matrix) model captures the current testing practice, as responding to engineering change during product development is particularly dependent on understanding how testing activities (which often drive and validate the changes) are integrated with design activities as well as how the testing activities themselves are related. This ADSM essentially offers the potential to restructure the testing activities in the PD process. At the same time, the connectivity network between the testing and the components (i.e. DMM) can potentially be used for identifying how the propagation of design changes affects component and system testing, also helps to visualise how effectively testing information are used.

The third question investigates a way to improve the overlapping between testing and design activities:

> How can overlapping be effective, even where the upstream evolution of information is slow and downstream sensitivity is high?

In the literature review in Chapter 2, it is described that overlapping is a technique that is used to reduce the development time. But in both case study companies, it has been found that engineers do not plan to overlap stages of the PD. Testing activities often cannot be finished within the gateway stages and spill over to the next stage. After the gateway review, engineers have no choice but to start downstream design overlapping with upstream testing to meet the deadline for the next stage. Such overlapping is not deliberately planned but commonly happens in product development projects.

This research has studied the issues that cause these unintentional instances of overlapping in the process. A method is proposed to reduce the uncertainty caused by these overlapping by using intermediary testing measurements, rather than just using tacit knowledge and overview of experienced engineers. This method also suggests using virtual testing in parallel with physical testing to mirror physical testing measurements. The proposition of this method is that if virtual models are accurately built and validated through sufficient and necessary intermediary testing measurements, simulation of those virtual models will be able to predict, to some measure of accuracy, future testing results, faster than the actual testing. Thus, these simulated results will be adequate to start downstream design; such that design will not suffer considerable rework. This method is validated though discussions with senior engineers in both companies. An example was created to demonstrate the potential application of this method. The effect is measured by comparing the rework time that will require if 'this method is applied' or 'not applied'. Engineers have realised that this method potentially can improve the current process by reducing uncertainties in testing results and reduce the rework time in subsequent design.

The fourth question explores a way of integrating virtual and physical testing:

> How can virtual and physical testing results be integrated so that the overall capability of physical testing can be improved?

The literature and case study have suggested that while both virtual and physical test have their own advantages and limitations, a combined approach of physical and virtual testing might help to produce a focused test and increase reliability in testing. Although, the case study company has an established CAE team and skilled teams of design and testing engineers, there are still deficiencies in the virtual testing models. One reason was identified is that these models may be built upon previous physical testing results and are used for simulating the future physical testing. As the design changes, the testing conditions and properties significantly change also. This research suggests to calibrate and validate the virtual models with instantaneous testing results. Essentially, using the same method explained in third question, above, i.e. the method of concurrent execution of virtual and physical testing, in which, the virtual model is validated with sufficient and necessary testing measurements during a physical test. When virtual and physical testing results converge and produce the similar results, engineers can achieve higher confidence in their decisions. This improves the capability of virtual testing; later this virtual model is simulated to predict the testing results. This parallel integration method proves to be having better capabilities than the iterative execution of virtual and physical testing.

The fifth question concerns the structural changes of the PD process:

> How does the structure of the product development process changes with the integration of virtual testing into the stages of product development?

This study has identified that virtual testing is a form of CAE but there is a difference between 'initial CAE analysis of design' and 'virtual testing' (see Chapter 5 for details). These differences are based on the 'time of use' and 'intension of use' of these CAEs. CAE analysis is used to analyse the design and virtual testing is used to test the intended design. Virtual testing is the advanced and focused form of CAE modelling and simulation that can be a continuation of CAE analysis. Usually design information is released to other departments, especially, for procurement, after initial CAE analysis. Following these stages the design can be tested through both virtual and physical testing.

Both CAE and virtual testing are performed at each phase of the product development process. The case study company's current product development structure does not reflect these differences and presents CAE analysis and virtual testing together as CAE (can be seen in Chapter 5 (Figure 5.8). A model of the product development process is proposed that separates virtual testing from the initial CAE analysis and positions this virtual testing in such a way that the role of these computational activities can be realised (the schematics of the proposed product development structure can be found in Figure 8.18). This study proposes structural changes of the company's product development process by introducing virtual testing in parallel to the physical testing in each product development phase.

9.1.2 Research findings and contributions

With emerging or new technologies entering the market through new or existing products, thorough testing before release is critical. Thus the demands for testing are increasing significantly and conditions are getting tougher in terms of testing scope, coverage, and rigor. These are set to increase, because of pressures to improve quality. However, testing is relatively under-researched in comparison to other fields of design. As, described in the literature review, Chapter 2, there are relatively few publications on this topic. Most discuss testing in terms of specific techniques. Little analytical effort has been made to understand the role of testing and efficiently plan testing during the product development. This thesis has addressed this by reporting upon an in-depth empirical study carried out in a UK based engineering industry. This thesis contains a comprehensive literature review on testing-related terms and activities, different modes of testing, and how testing might be integrated into the product development process. At a general level, the work described in this thesis has developed the understanding of the role of testing in the engineering product development process. This is probably one of the first serious studies that investigates the role of testing and analyses its implication for the product development process.

9.1.2.1 Empirical findings

- Testing is major driver in engineering product design and development process.
- In companies, testing starts early at the conceptual design phase, even before the detailed design starts and continues almost in parallel with design.
- By analysing influencing factors, objectives and reasons of testing, it can be identified the critical testing activities that eventually provides useful inputs for testing planning (Section 4.4-4.5).
- Four entities that characterise the testing are: the property that is tested, the mode of testing, object under test and location of the test. Testing planning is a process of finding a balance of these entities (Section 5.1).
- People involved in these planning aspects focus on different factors, such as, need, risk, cost, time. Therefore, they might not always share the same understanding of the relative importance of these factors.
- Test planning involves active processes, starting at the beginning of product development but constantly modified and adjusted during the process; because design changes introduce and/or eradicate testing and testing introduces design changes.

- Most engineers know that the process they are following may not be the best one and there can be better alternatives, but they are bound to follow the organisational practices in the company or group, often forced to react on serious problems (fire-fighting approach). Although they realise the benefit, engineers who often cannot spend time to analyse and learn from the mistakes of previous projects. Lack of analysis can foster inefficiency in the process and cause engineers to repeat mistakes.
- Both computer aided analysis and computer aided testing are types of CAE (computer aided engineering) but play different roles in different times of the PD process. Therefore, it is required to identify the role of these CAEs and distinguish these so that associated activities can be planned and performed accurately.
- Long lead time for procurement and lengthy physical testing were identified as factors that force the testing and design tasks to overlap, therefore causing uncertainties in the process.

9.1.2.2 Theory building

- Testing is not a standalone phase towards the end of the product development process but starts from the beginning to assure the quality of design throughout the development process. Consequently testing is closely intertwined with design.
- The roles of testing change in different stages of product development, these roles need to be clearly identified so that these testing activities can be planned accurately in different stages of product development.
- If a virtual model is accurately validated with practical intermediate physical testing data and used appropriately, simulation of the model will predict the test results more precisely.
- Virtual testing can minimise the issues arise during overlapping of testing and design tasks, if accurately validated with practical intermediate physical testing data and used appropriately.

9.1.2.3 Towards a tool

- A method of prioritisation of testing activities, based on critical customer requirements and critical products characteristics is suggested, so that critical tests can be recognised and planned accurately (presented in Chapter 6).

- A conceptual method of integrating virtual and physical testing is proposed. In which, the virtual model is validated and calibrated with intermediary physical test measurements, before this model can be used to simulate the test results and transferred to the downstream tasks in a preliminary form, in the case of overlapping (presented in Chapter 8).
- This research suggests a modification to the existing product development process structure for designing and testing, which explicitly recognises virtual testing as a distinct and significant activity to assist physical testing (presented in Chapter 8).

9.1.2.4 Contributions to knowledge

- An extensive industrial case study was performed, which provided new insights from industrial practice and better understanding of the role of testing in the product development process.
- This case study provides evidence that testing and design are intertwined throughout the design process.
- This research identifies the causes of overlap between testing and design activities; this unintentional overlapping is a source of iteration and inefficiency in the process. Also develops conceptual methods for addressing these issues of overlapping.
- This research develops ideas and models for improving the test planning considering the interrelation between design and testing in different stages of product development process by gaining the perspectives of different stakeholders.
- This study claims that the conventional product development process structures require modification to highlight the role of different activities during the process more clearly.

9.1.2.5 Contribution to the case study company

- Engineers obtained external views and different insights on issues in the process.
- This research brought academic findings into the company.
- Through the discussions during the interviews, people from different departments shared their views and opinions on issues.
- Methods and procedures were discussed with the engineers, which initiated some 'internal thought processes' on changes required in the business.

9.2 Research validation

Given the nature of this study and the methods used, rigorous scientific evaluation procedures may not be applicable for the methods, tools and process improvements which are proposed. However, they have been carefully considered where appropriate to widen the acceptance of this research.

New methods and tools that have been developed require evaluation in order to establish their validity. However, validation of design methods and tools can be a difficult task because evaluation is very much dependent on the context in which the tools and methods are applied. Each design process presents a unique context driven by the customer, product and company.

This section describing the validation of the research outcomes is divided into three parts: (a) empirical findings, (b) theory building and (c) towards tools. These outcomes are validated using a validation square, which was presented as part of the research methodology in Section 3.3. The overall scheme of validation is laid out in Figure 9.1 which also highlights the main areas of contribution presented in the previous section.

The initial supposition of this research was that '*testing should be integrated in the product development process and is not a standalone phase to be performed towards the end of the processes*'. The *theoretical structural validity* of the initial supposition was achieved through extensive literature review in Chapter 2. Each of the concepts was considered individually and the logical connections among these concepts established.

The *empirical structural validity* – has been achieved by accepting that the case of diesel engine as an appropriate example problem. The case is valid because it presents the characteristics of a complex engineering product and process, for which the initial propositions are accepted.

In the assessment of *empirical performance validity* the core issue considered was the extent to which the outcome of the research is useful with respect to the initial purpose particularly for an example problem. One indicator of the usefulness of empirical findings is the successful production more scientific knowledge and its associated peer review from an academic perspective. The findings were presented in academic papers, have been subject to full reviews before acceptance. Through the published papers' acceptance by other researchers, value and contribution have been established.

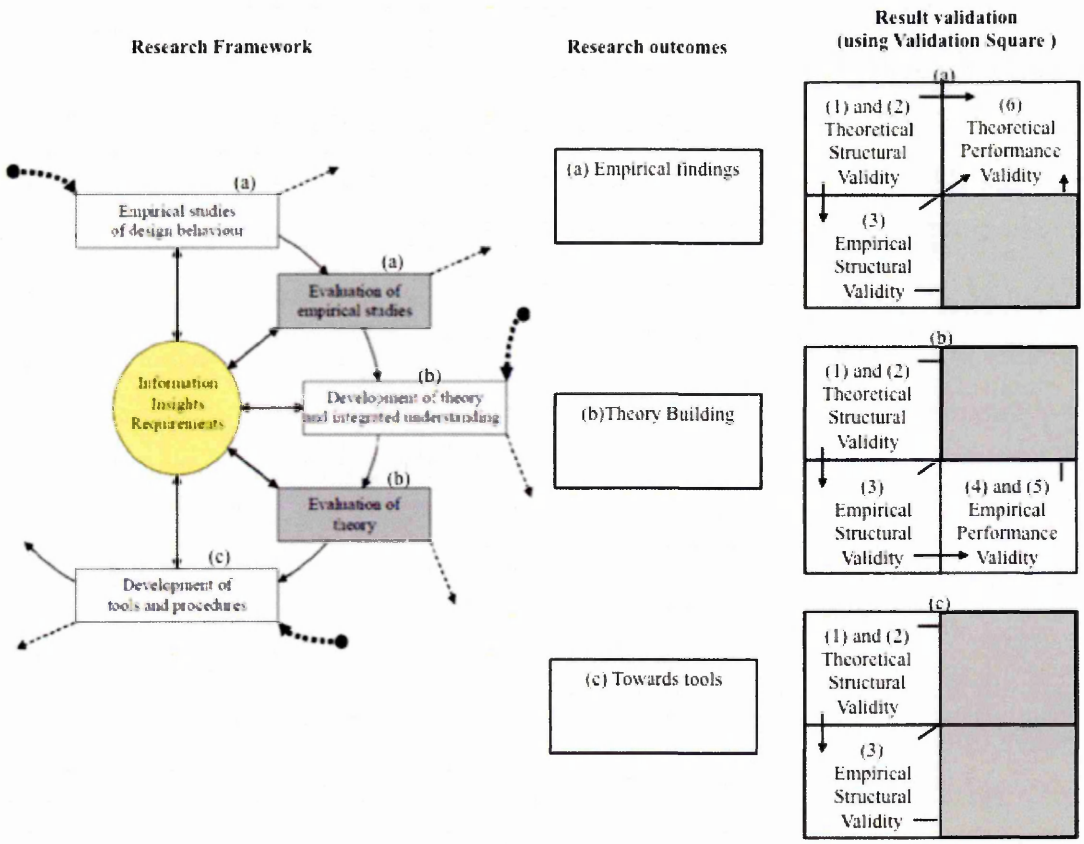


Figure 9.1 Research framework, outcomes and validation

The *theoretical performance validity* can be demonstrated through showing the usefulness of the research beyond the example problems. Generalisation is attained through a small set of interviews in a different company as well as being confirmed in informal discussions with academics and company engineers.

9.3 Limitations of this work

Alongside the contributions described above, the research presented here will have its limitations. This section attempts to cover some of these possible limitations.

One of limitation is that findings are based mainly on a single case study. Although, another study was conducted in a different company it was in a limited form. Analysis of additional industrial products and processes would be required to confirm the generality of research findings across multiple contexts.

This research identified a subtle difference between CAE analysis and virtual testing but could not provide a definitive, firm boundary between these activities which would differentiate them clearly. As virtual testing can be a progression from CAE analysis, it is not clear when exactly a CAE analysis becomes a process of virtual testing. The thesis did present a provisional definition based on the stage in the process, namely that CAE analysis after design sign off was considered as virtual testing. Confirming the details of such a definition as workable and applicable across design contexts would require wider observation of uses and detailed analysis of CAE tools and virtual models. Overall, a more thorough understanding of this progression between CAE analysis and virtual testing in different contexts would be needed to define the boundary.

Simulations of proposed process models could be useful to estimate the effect of integrating virtual testing with physical testing and thus validate this part of the thesis. Although a group of testing activities were modelled and simulated to estimate the time savings, it was not possible to model the overall product development process. To simulate the overall process, a detailed process model with an estimation of testing time and cost would be required. That can only be provided by experienced engineers who have adequate knowledge of planning the validation and testing activities. Although, several attempts have been made by the researcher to obtain this information, the engineers in the case study company were unable to commit the required time for this. To estimate the cost of integrating virtual testing with physical testing, a model was developed. However, it was not possible to evaluate its use in a real context. Again, this would need engineers' time and effort from the company but that was not possible to attain.

One further point of note is that although the Cambridge Advanced modeller (CAM) tool was used for process modelling, this tool yet does not have the capability of concurrent process modelling. The time to explore different tools and their application was not available.

9.4 Future work

In general terms, the future work will focus upon developing the proposed methods further and examining their wider applicability across design processes, as well as their use for the other potential applications in different industry examples.

Research which is an immediate extension of the thesis could model and simulate the proposed product development structure that integrates the virtual testing in parallel to physical testing. The first objective would be to demonstrate reduction in uncertainties in overlapping in the original case study company. Process modelling and simulation will evaluate the benefit of using this proposed structure. Further work will extend validation of this model in other industrial contexts. In particular, overlapping considerations for the design and testing of products at different scale, complexity and maturity could be compared. Further the model could be extended to consider multiple layered overlapping.

Chapter 8 proposed a conceptual method of integrating virtual and physical testing. This method can be further developed by including other aspects, such as:

(i) using intermediary testing measurements taken from a physical tests;
(ii) understanding the format of available information and data;
(iii) assessing the extent to which data and information can be readily used for virtual models;
(iv) processing data (if such data is not in a form that can be readily used) ready for the CAE team to use;
(v) comparing the physical test measurements and virtual simulated results;
(vi) aggregating virtual and physical test results.
(vii) assessing the required time, effort and resources required for the above tasks.

Future work will further develop the understanding of the difference between CAE analysis and virtual testing. Detailed studies will be required with not only those who use the CAE tools but also with those who develop them. These studies will serve to help identify the current capabilities of CAE tools and the specialised bespoke use of these tools for the purpose of virtual testing. Also, detailed observations how these tools are actually used will help to pinpoint the transitions between CAE analysis and virtual testing.

Chapter 4 of this thesis describes how the case study company performs failure mode and effect analysis (FMEA) and assesses the risks of a product development programme. Currently, risk analysis is a top-down approach, i.e. product level risks are broken down into sub-system and component level risks. Risks mitigated into the lower levels are counted as mitigation of top levels risks. Although, each component may be proven to be fail-safe, often top-level failures are caused by multiple components functioning together. The current approaches of mitigating each lower level individual risks does not ensure that the top level risks have been mitigated. Therefore, an approach of calculating top level risk in a bottom-up approach will be developed that could improve the overall technical risk assessment of a product.

Chapter 6 of this thesis describes how prioritisation of testing activities can be improved considering both risk and importance of technical objectives. This method suggests using data from QFD and FMEA analysis. Currently, these QFD, FMEA tools are separated and managed by different groups of people. Future work could develop a software tool, which will integrate these QFD and FMEA approach in a single platform. This approach use the prioritisation method proposed in Chapter 6 for automatically ranking testing activities.

In the longer term, a significant effort could be made to revise general models of design to give testing the prominence which has been identified in this thesis. In particular, effort might usefully be directed to integrate design, analysis and virtual testing in a single platform. At each stage of product development process, from concept demonstration to product validation, different CAE analysis and modelling technology exists. But there does not exist any methodological approach of modelling and simulating that a company can use to meet the requirements in each stage. Hence, there is a need for a scalable approach, involving the combination of multiple and often-heterogeneous simulation methods into full system models (Van der Auweraer et al. 2013). There will also be a need for a single platform that will provide connectivity between design, simulation tools and virtual testing. There will be a strong need for an easy to use, modular, customisable but open COTS (commercial off-the-shelf) test software platform with a plug-and-play architecture (Truchard 2005).

For implementing these, companies will need investment for these technological advances. Despite the fact that there are requirements to implement new and effective technologies and acquire technical skills, many companies are reluctant to move towards major investment especially, automated, intelligent and Computer Aided Engineering systems. The reason is that the capabilities of the technology, especially in simulation and virtual testing, and the advantages it brings may not be fully understood in many companies. Moreover, physical testing is a necessary industrial practice, usually required for product certification. For example, aerospace industries undertake a rigorous testing regime to pass certification criteria and automobile manufacturers test their prototypes for regulatory and safety standards (Maropoulos & Ceglarek 2010). So an industrial shift is needed towards greater acceptance of computer aided, virtual testing among companies and certification bodies as part of their product development processes.

References

AKAO. Y.,,2004, *Quality function deployment.* Taylor & Francis.

AL KINDI, M. and ABBAS, A., 2010. Sequential testing decisions in product development, *Systems Conference, 2010 4th Annual IEEE* 2010, IEEE, pp. 490-494.

ALLEN, L., ANGEL, J., MONGUS, J., RODNEY, G.A., SHANNON, R.R. and SPOELHOF, C.P., 1990. The Hubble Space Telescope optical systems failure report. *NASA Report.*

AMMANN, P. and OFFUTT, J., 2008. *Introduction to software testing.* Cambridge Univ Pr.

AUSTIN, S., STEELE, J., MACMILLAN, S., KIRBY, P. and SPENCE, R., 2001. Mapping the conceptual design activity of interdisciplinary teams. *Design Studies,* **22**(3), pp. 211-232.

AVEN, T. and ZIO, E., 2011. Some considerations on the treatment of uncertainties in risk assessment for practical decision making. *Reliability Engineering & System Safety,* **96**(1), pp. 64-74.

AYAG, Z., 2005. An integrated approach to evaluating conceptual design alternatives in a new product development environment. *International Journal of Production Research,* **43**(4), pp. 687-713.

BALCI, O., 2003. Verification, validation, and certification of modeling and simulation applications: verification, validation, and certification of modeling and simulation applications, *Proceedings of the 35th conference on Winter simulation: driving innovation* 2003, Winter Simulation Conference, pp. 150-158.

BALCI, O., 1998. Verification, validation, and testing. *Handbook of simulation*, pp. 335-393.

BALDWIN, C.Y. and CLARK, K.B., 2000. *Design rules: The power of modularity.* The MIT Press.

BARNARD, R.W.A. and CONSULTING, L., 2008. Integration of Reliability Engineering into Product Development, *2nd South African International Aerospace Symposium* 2008, pp. 16.

BARRIENTOS, F.A., TUMER, I.Y. and ULLMAN, D.G., 2010. Design teams, complex systems and uncertainty. *Relation,* **10**(1.129), pp. 2496.

BARTOLOMEI, J.E., HASTINGS, D.E., DE NEUFVILLE, R. and RHODES, D.H., 2012. Engineering Systems Multiple-Domain Matrix: An organizing framework for modeling large-scale complex systems. *Systems Engineering,* **15**(1), pp. 41-61.

BASILI, V.R. and SELBY, R.W., 1987. Comparing the effectiveness of software testing strategies. *Software Engineering, IEEE Transactions on,* (12), pp. 1278-1296.

BECK, K., 2000. *Extreme programming explained: embrace change.* Addison-Wesley Professional.

BECK, K., 1999. Embracing change with extreme programming. *Computer,* **32**(10), pp. 70-77.

BECKER, M.C., SALVATORE, P. and ZIRPOLI, F., 2005. The impact of virtual simulation tools on problem-solving and new product development organization. *Research Policy,* **34**(9), pp. 1305-1321.

BERTOLINO, A., 2009. Software Testing Research: Achievements, Challenges, Dreams. *IEEE Transactions on Software Engineering,* **35**(4).

BOEHM, B. and HANSEN, W.J., 1998. Spiral Development: Experience, Principles, and Refinements. (No CMU/SEI-2000-SR-008) Carnegie-Mellon University Pittsburgh PA Software Engineering Institute.

BOGUS, S.M., MOLENAAR, K.R. and DIEKMANN, J.E., 2006. Strategies for overlapping dependent design activities. *Construction Management and Economics,* **24**(8), pp. 829-837.

BOUCHER, M., 2010. *Cost Saving Strategies for Engineering: Using Simulation to Make Better Decisions* Boston, MA: Aberdeen group.

BRINGMANN, E., 2008. Model-Based Testing of Automotive Systems, *International Conference on Software Testing, Verification, and Validation* 2008, IEEE, pp. 485-493.

BROWNING, T.R., 2001. Applying the design structure matrix to system decomposition and integration problems: a review and new directions. *Engineering Management, IEEE Transactions on,* **48**(3), pp. 292-306.

BROWNING, T.R., FRICKE, E. and NEGELE, H., 2006. Key concepts in modeling product development processes. *Systems Engineering,* **9**(2), pp. 104-128.

CHAKRAVARTY, A.K., 2001. Overlapping design and build cycles in product development. *European Journal of Operational Research,* **134**(2), pp. 392-424.

CHEN, S.J.G. and LIN, L., 2003. Decomposition of interdependent task group for concurrent engineering* 1. *Computers & Industrial Engineering,* **44**(3), pp. 435-459.

CHILES, J.R., 2002. Inviting disaster: Lessons from the edge of technology: An inside look at catastrophes and why they happen. *Inviting disaster: lessons from the edge of technology: an inside look at catastrophes and why they happen/James R.Chiles.New York: Harper Business, 2002.,* **1**.

CHISHOLM, I., 1984. Automobile Engine and Vehicle Technology: Level 1. McGraw-Hill.

CHUA, C., TEH, S. and GAY, R.K.L., 1999. Rapid prototyping versus virtual prototyping in product design and manufacturing. *The International Journal of Advanced Manufacturing Technology,* **15**(8), pp. 597-603.

CLARK, K.B. and FUJIMOTO, T., 1987. *Overlapping problem solving in product development.* Division of Research, Harvard Business School.

CLARKSON, J. and ECKERT, C., 2005. *Design process improvement: a review of current practice.* Springer.

CLARKSON, P.J., SIMONS, C. and ECKERT, C., 2004. Predicting change propagation in complex design. *Journal of Mechanical Design,* **126**, pp. 788.

CLARKSON, P.J. and HAMILTON, J.R., 2000. 'Signposting', a parameter-driven task-based model of the design process. *Research in Engineering Design,* **12**(1), pp. 18-38.

COOPER, R.G., 1990. Stage-gate systems: a new tool for managing new products. *Business horizons,* **33**(3), pp. 44-54.

CRESWELL, J.W., 2009. *Research design: Qualitative, quantitative, and mixed methods approaches.* Sage Publications, Inc.

DAHAN, E. and MENDELSON, H., 2001. An extreme-value model of concept testing. *Management Science,* , pp. 102-116.

DASSO, A. and FUNES, A., 2007. *Verification, validation and testing in software engineering.* IGI Global.

DFKI, 2011-last update, DVP&R - Design Verification Plan and Report. Available: http://standardsengineer.com/Glossary/Glossary_d/DVP&R.htm [22/04, 2014].

DONG, Q., 2002. *Predicting and Managing System Interactions at early phase of the product development process*, Massachusetts Institute of Technology.

EARL, C., ECKERT, C.M. and JOHNSON, J., 2000. *Complexity of planning in design.* Cambridge Engineering Design Centre.

ECKERT, C., CLARKSON, J., DE WECK, O. and KELLER, R., 2009. Engineering Change: Drivers, Sources, and Approaches in Industry, *Proceedings of the 17th International Conference on Engineering Design (ICED'09), Vol. 4* 2009, pp. 47-58.

ECKERT, C.M., STACEY, M.K. and CLARKSON, P.J., 2003. The spiral of applied research: a methodological view on integrated design research.

ENGEL, A., 2010. *Verification, validation, and testing of engineered systems.* Wiley.com.

EPA'S OFFICE, 23/01/2013, 2013-last update, Nonroad Engines, Equipment, and Vehicles. Available: http://www.epa.gov/nonroad/ [11/03, 2014].

EPPINGER, S.D. and BROWNING, T.R., 2012. *Design structure matrix methods and applications.* MIT Press (MA).

EPPINGER, S.D., WHITNEY, D.E. and GEBALA, D.A., 1992. Organizing the tasks in complex design projects: Development of tools to represent design procedures, *Proceedings NSF Design and Manufacturing System Conference, Atlanta, Georgia* 1992.

ERAT, S. and KAVADIAS, S., 2008. Sequential testing of product designs: Implications for learning. *Management Science,* **54**(5), pp. 956-968.

FERRY, W., FRISE, P., ANDREWS, G. and MALIK, M., 2002. Combining virtual simulation and physical vehicle test data to optimize durability testing. *Fatigue & Fracture of Engineering Materials & Structures,* **25**(12), pp. 1127-1134.

GERK, J.E.V. and QASSIM, R.Y., 2008. Project acceleration via activity crashing, overlapping, and substitution. *Engineering Management, IEEE Transactions on,* **55**(4), pp. 590-601.

HANKS, M.L., 2009. Failures Due to Product Changes: The Importance of Performance Testing when Making Product Modifications. *Journal of failure analysis and prevention,* **9**(6), pp. 499-503.

HAUSER, J.R. and CLAUSING, D., 1988. The house of quality. *Harvard Business Review, 66*(3), 63-73.

HELLE, P. and SCHAMAI, W., 2013. Towards an Integrated Methodology for the Development and Testing of Complex Systems, *VALID 2013, The Fifth International Conference on Advances in System Testing and Validation Lifecycle* 2013, pp. 55-60.

HELMREICH, K. and REINWARDT, G., 1996. Virtual test of noise and jitter parameters, *Test Conference, 1996. Proceedings., International* 1996, IEEE, pp. 461-470.

HOLMES, R., ROBERT, G.T.I., POPE, J.E., OJA, M., DEY, A., HAMLER, J., BROOKS, J., TRYON, J., NASSER, D.A. and NASSER, A.L., 2011. *Testing performance of a material for use in a jet engine.* Google Patents.

HOPPE, M., ENGEL, A. and SHACHAR, S., 2007. SysTest: Improving the verification, validation, and testing process—Assessing six industrial pilot projects. *Systems Engineering,* **10**(4), pp. 323-347.

HOPPE, M., LÉVÁRDY, V., LEARDI, C., MENDIKOA, I., DE ABAJO, N., GONZÁLEZ, J.A., LOBATO, V. and PEREGRINA, S., 2004. Application Experiences of the VVT Process Modeling Procedure for Verification and Validation Planning, *Proceedings of the 14th Annual International Symposium of INCOSE* 2004, pp. 24.

HOY, T.W., CENTER, U.S.A.A.T. and HERRMANN, J.W., 2008. Optimal Utilization of Test Facilities to Replicate Operational Environments. *Journal of the IEST,* **51**(2), pp. 10.

HOYLE, D., 2006. *ISO 9000 quality systems handbook.* Butterworth-Heinemann.

HUIZINGA, F., VAN OSTAIJEN, R.A.A. and SLINGELAND, A.V.O., 2002. A practical approach to virtual testing in automotive engineering. *Journal of Engineering Design,* **13**(1), pp. 33-47.

ISHIZAKA, A. and LABIB, A., 2011. Review of the main developments in the analytic hierarchy process. *Expert Systems with Applications,* **38**(11), pp. 14336-14345.

JARRATT, T.A., 2004. *A model-based approach to support the management of engineering change*, Cambridge University.

JONES, H.W., 2007. Integrated Systems Testing of Spacecraft. (No. 2007-01-3144). SAE Technical Paper.

JONES, J., 2/03/2011, 2011-last update, Reducing the time and cost of testing engines [Homepage of NASA], [Online]. Available: http://www.sti.nasa.gov/tto/Spinoff2004/t_2.html [07/2011, 2011].

JONES, R. and SIMON, D.L., 5/9/02, 2002-last update, Spotlight Innovation, Not Just Evaluation. Available: http://www.altair.com/%28S%28ud4svw45pku4sh55m0gxdvfd%29%29/ResLibSearchResult.aspx?keywords=Spotlight%20Innovation,%20Not%20Just%20Evaluation [09/11, 2013].

KELLER, R., ECKERT, C. and CLARKSON, P., 2009. Using an engineering change methodology to support conceptual design. *Journal of Engineering Design,* **20**(6), pp. 571-587.

KHALAF, F., 2006a. An engineered testing strategy: part I: enhanced balance between analytic and hardware. *International Journal of Product Development,* **3**(3), pp. 404-418.

KHALAF, F., 2006b. An engineered testing strategy: part II: axiomatic design approach for balanced strategy. *International Journal of Product Development,* **3**(3), pp. 419-431.

KRISHNAN, V., EPPINGER, S.D. and WHITNEY, D.E., 1997. A model-based framework to overlap product development activities. *Management science,* , pp. 437-451.

KRISHNAN, V. and BHATTACHARYA, S., 2002. Technology selection and commitment in new product development: The role of uncertainty and design flexibility. *Management Science,* **48**(3), pp. 313-327.

LAWLER, E.L., 2001. *Combinatorial optimization: networks and matroids.* Courier Dover Publications.

LEE, D.C. and HAN, C.S., 2009. CAE (computer aided engineering) driven durability model verification for the automotive structure development. *Finite Elements in Analysis and Design,* **45**(5), pp. 324-332.

LÉVÁRDY, V. and BROWNING, T.R., 2009. An adaptive process model to support product development project management. *Engineering Management, IEEE Transactions on,* **56**(4), pp. 600-620.

LÉVÁRDY, V., HOPPE, M. and BROWNING, T.R., 2004. Adaptive Test Process: An Integrated Modeling Approach for Test and Design Activities in the Product Development Process, 2004, ASME.

LIN, J., CHAI, K.H., WONG, Y.S. and BROMBACHER, A.C., 2008. A dynamic model for managing overlapped iterative product development. *European Journal of Operational Research,* **185**(1), pp. 378-392.

LINDEMANN, U., 2012-last update, The Design Structure Matrix (DSM) [Homepage of Institute of Product Development Technische Universität München], [Online]. Available: http://www.dsmweb.org/ [11/2012, 2012].

LOCH, C.H., TERWIESCH, C. and THOMKE, S., 2001. Parallel and sequential testing of design alternatives. *Management Science,* , pp. 663-678.

LOCH, C.H., DEMEYER, A. and PICH, M.T., 2011. *Managing the unknown: A new approach to managing high uncertainty and risk in projects.* Wiley.

LU, P., GLASER, D., UYGUR, G. and HELMREICH, K., 2009. A novel approach to entirely integrate virtual test into test development flow, *Proceedings of the Conference on Design, Automation and Test in Europe* 2009, European Design and Automation Association, pp. 797-802.

MAROPOULOS, P.G. and CEGLAREK, D., 2010. Design verification and validation in product lifecycle. *CIRP Annals-Manufacturing Technology,* .

MARTIN, A.C.M. and DE CARVALHO, M.M., 2006. Applying virtual simulation in automotive new product development process. **4**(2), pp. 79.

MATHUR, A.P., 2008. *Foundations of software testing.* Pearson Education India.

MAURER, M. and LINDEMANN, U., 2008. The application of the Multiple-Domain Matrix: Considering multiple domains and dependency types in complex product design, *Systems, Man and Cybernetics, 2008. SMC 2008. IEEE International Conference on* 2008, IEEE, pp. 2487-2493.

MAXWELL, J.A., 1992. Understanding and validity in qualitative research. *Harvard educational review,* **62**(3), pp. 279-301.

MCMANUS, H. and HASTINGS, D., 2006. A framework for understanding uncertainty and its mitigation and exploitation in complex systems. *IEEE Engineering Management Review,* **34**(3), pp. 81-94.

MENOR, L.J., TATIKONDA, M.V. and SAMPSON, S.E., 2002. New service development: areas for exploitation and exploration. *Journal of Operations Management,* **20**(2), pp. 135-157.

MILLER, L. and MILLER, R., 2012. Classifying Innovation. *International Journal of Innovation and Technology Management,* **9**(01),.

MURTHY, D.P., RAUSAND, M. and ØSTERÅS, T., 2008. *Product reliability: specification and performance.* Springerverlag.

MYERS, G.J., SANDLER, C. and BADGETT, T., 2011. *The art of software testing.* John Wiley & Sons.

NELSON, W.B., 2009. *Accelerated testing: statistical models, test plans, and data analysis.* Wiley. com.

NOVAK, S. and EPPINGER, S.D., 2001. Sourcing by design: product complexity and the supply chain. *Management science,* **47**(1), pp. 189-204.

O'CONNOR, P.D.T., 2001. *Test engineering: a concise guide to cost-effective design, development and manufacture.* New York: Wiley.

O'CONNOR, P.D.T., NEWTON, D. and BROMLEY, R., 2002. *Practical reliability engineering.* John Wiley & Sons Inc.

OSTERAS, T., MURTHY, D. and RAUSAND, M., 2006. Product performance and specification in new product development. *Journal of Engineering Design,* **17**(2), pp. 177-192.

OULD, M.A. and UNWIN, C., 1986. *Testing in software development.* Cambridge University Press.

PAN, J., 1999. Software reliability. *Dependable Embedded Systems, Carnegie Mellon University,* .

PARK, T. and KIM, K., 1998. Determination of an optimal set of design requirements using house of quality. *Journal of Operations Management,* **16**(5), pp. 569-581.

PEDERSEN, K., EMBLEMSVAG, J., BAILEY, R., ALLEN, J.K. and MISTREE, F., 2000. Validating design methods and research: the validation square, *ASME Design Theory and Methodology Conference* 2000.

PERRY, W.E., 1999. *Software Testing.* John Wiley & Sons, Inc.

PINEDA, R.L. and KILICAY-ERGIN, N., 2010. System Verification, Validation, and Testing. *Systems Engineering Tools and Methods,* , pp. 81.

PRIDDLE, B., SINGLETON, I., LOO, M. and FOUNTAIN, M., 2003. BA Falcon Intelligent Safety Systems: A Systems Engineering Approach, 2003, Road Safety Research, Policing and Education Conference.

QIAN, Y., XIE, M., GOH, T.N. and LIN, J., 2010. Optimal testing strategies in overlapped design process. *European Journal of Operational Research,* **206**(1), pp. 131-143.

ROEMER, T.A., AHMADI, R. and WANG, R.H., 2000. Time-cost trade-offs in overlapped product development. *Operations research,* , pp. 858-865.

ROEMER, T.A. and AHMADI, R., 2004. Concurrent crashing and overlapping in product development. *Operations research,* **52**(4), pp. 606-622.

ROSS, A.M., HASTINGS, D.E., WARMKESSEL, J.M. and DILLER, N.P., 2004. Multi-attribute tradespace exploration as front end for effective space system design. *Journal of Spacecraft and Rockets,* **41**(1), pp. 20-28.

SECRETARIAT, E., 2002. *Spacecraft Engineering Testing.* ESCC.

SEGISMUNDO, A. and MIGUEL, P.A.C., 2008. Failure mode and effects analysis (FMEA) in the context of risk management in new product development: A case study in an automotive company. *International Journal of Quality & Reliability Management,* **25**(9), pp. 899-912.

SEPTEMBER, A., 1990. IEEE Standard Glossary of Software Engineering Terminology. *Office,* **121990**(1),.

SHABI, J. and REICH, Y., 2012. Developing an analytical model for planning systems verification, validation and testing processes. *Advanced Engineering Informatics,* .

SHAHIN, A., POURHAMIDI, M., ANTONY, J. and PARK, S.H., 2013. Typology of Kano models: a critical review of literature and proposition of a revised model. *International Journal of Quality & Reliability Management,* **30**(3), pp. 341-358.

SHAMSUZZOHA, A.H.M. and HELO, P.T., 2011. Information dependencies within product architecture: prospects of complexity reduction. *Journal of Manufacturing Technology Management,* **22**(3), pp. 314-329.

SIEMENS, 2014-last update, CAE / Computer-Aided Engineering. Available: http://www.plm.automation.siemens.com/en_gb/plm/cae.shtml [05/20, 2014].

SONG, X.M. and MONTOYA-WEISS, M.M., 1998. Critical development activities for really new versus incremental products. *Journal of Product Innovation Management,* **15**(2), pp. 124-135.

SOSA, M.E., EPPINGER, S.D. and ROWLES, C.M., 2003. Identifying modular and integrative systems and their impact on design team interactions. *Journal of Mechanical Design,* **125**, pp. 240.

STAMATIS, D.H., 2003. *Failure mode and effect analysis: FMEA from theory to execution.* Asq Press.

STEWARD, D.V., 1981. Design structure system: A method for managing the design of complex systems. *IEEE TRANS.ENG.MGMT.,* **28**(3), pp. 71-74.

SUH, N.P., 2001. Axiomatic Design: Advances and Applications (The Oxford Series on Advanced Manufacturing).

TAHERA, K., EARL, C. and ECKERT, C., 2012. The role of testing in the engineering product development process. Proceedings of TMCE 2012, 7-11 May 2012, Karlsruhe, Germany.

TAHERA, K., ECKERT, C.M. and EARL, C.F., 2013. Improving overlapping between testing and design in engineering product development processes. In: ASME 2013 International Design Engineering Technical Conferences (IDETC) and Computers and Information in Engineering Conference (CIE) , 4-7 August 2013 , Portland, Oregon, USA.

TATIKONDA, M.V. and ROSENTHAL, S.R., 2000. Technology novelty, project complexity, and product development project execution success: a deeper look at task uncertainty in product innovation. *Engineering Management, IEEE Transactions on,* **47**(1), pp. 74-87.

TERWIESCH, C. and LOCH, C.H., 1999. Measuring the effectiveness of overlapping development activities. *Management Science,* , pp. 455-465.

TERWIESCH, C., LOCH, C.H. and DE MEYER, A., 2002. Exchanging preliminary information in concurrent engineering: Alternative coordination strategies. *Organization Science,* , pp. 402-419.

THOMKE, S. and BELL, D.E., 2001. Sequential testing in product development. *Management Science,* , pp. 308-323.

THOMKE, S.H., 2003. *Experimentation matters: unlocking the potential of new technologies for innovation.* Harvard Business Press.

THOMKE, S.H., 1998. Simulation, learning and R&D performance: Evidence from automotive development. *Research Policy,* **27**(1), pp. 55-74.

THOMKE, S. and FUJIMOTO, T., 2000. The Effect of "Front-Loading" Problem-Solving on Product Development Performance. *Journal of Product Innovation Management,* **17**(2), pp. 128-142.

THOMKE, S. and REINERTSEN, D., 2012. Six myths of product development. *Harvard business review,* **90**(5), pp. 84-94.

THOMKE, S., 2001. Enlightened experimentation. The new imperative for innovation. *Harvard business review,* **79**(2), pp. 66-75.

TRUCHARD, J., 2005. The emerging role of physical test in product development. *Automotive Engineering International,* **113**(10), pp. 90.

ULLMAN, D.G., 1997. *The mechanical design process.* New York: McGraw-Hill.

ULRICH, K.T. and EPPINGER, S.D., 2000. *Product design and development.* McGraw-Hill.

ULRICH, K.T. and EPPINGER, S.D., 2003. *Product design and development.* Tata McGraw-Hill Education,

ULRICH, K., 1995. The role of product architecture in the manufacturing firm. *Research policy,* **24**(3), pp. 419-440.

UNGER, D. and EPPINGER, S., 2011. Improving product development process design: a method for managing information flows, risks, and iterations. *Journal of Engineering Design,* **22**(10), pp. 689-699.

VAN DER AUWERAER, H. and LEURIDAN, J., 2005. A New Testing Paradigm for Today's Product Development Process-Part 1. *Sound and Vibration,* **39**(9), pp. 14.

VAN DER AUWERAER, H., ANTHONIS, J., DE BRUYNE, S. and LEURIDAN, J., 2013. Virtual engineering at work: the challenges for designing mechatronic products. *Engineering with Computers,* , pp. 1-20.

VERYZER, R.W., 1998. Discontinuous innovation and the new product development process. *Journal of Product Innovation Management,* **15**(4), pp. 304-321.

WANG, J. and LIN, Y.I., 2009. An overlapping process model to assess schedule risk for new product development. *Computers & Industrial Engineering,* **57**(2), pp. 460-474.

WASSERMAN, G., 2002. Reliability Verification Testing and Analysis in Engineering Design, Vol. 153 CRC Press.

WHEELWRIGHT, S.C., 2010. *Managing New Product and Process Development: Text Cases.* Free Press.

WILKINSON, P., 2007. The changing role of physical testing in vehicle development programmes. *Journal of Terramechanics,* **44**(1), pp. 15-22.

WYATT, D.F., ECKERT, C.M. and CLARKSON, P.J., 2009. Design of Product Architectures in Incrementally Developed Complex Products, *Proceedings of the 17th International Conference on Engineering Design (ICED'09), Vol. 4* 2009, pp. 167-178.

WYNN, D., 3/4/2011, 2011-last update, Cambridge Advanced Modeller. Available: http://www-edc.eng.cam.ac.uk/cam/ [7/12, 2011].

WYNN, D.C., 02/02/2013, 2013-last update, Simulating a process DSM to study concurrency and iteration [Homepage of Cambridge University], [Online]. Available: https://www-edc.eng.cam.ac.uk/cam/documentation/Using%20CAM%20to%20simulate%20a%20process%20DSM?action=view [04/2014, 2014].

WYNN, D.C., GREBICI, K. and CLARKSON, P.J., 2011. Modelling the evolution of uncertainty levels during design. *International Journal on Interactive Design and Manufacturing (IJIDeM),* **5**(3), pp. 187-202.

WYNN, D. and CLARKSON, J., 2005. Models of designing. *Design process improvement.* Springer, pp. 34-59.

YADAV, O.P., SINGH, N. and GOEL, P.S., 2006. Reliability demonstration test planning: A three dimensional consideration. *Reliability Engineering & System Safety,* **91**(8), pp. 882-893.

YANG, Q., LU, T., YAO, T. and ZHANG, B., 2013. The impact of uncertainty and ambiguity related to iteration and overlapping on schedule of product development projects. *International Journal of Project Management,* .

YASSINE, A. and BRAHA, D., 2003. Complex concurrent engineering and the design structure matrix method. *Concurrent Engineering,* **11**(3), pp. 165.

YASSINE, A, SREENIVAS, R.S. and ZHU, J., 2008. Managing the exchange of information in product development. *European Journal of Operational Research,* **184**(1), pp. 311-326.

YIN, R.K., 2009. *Case study research: Design and methods.* Sage Publications, Inc.

ZAKARIAN, A., 2010. A methodology for the performance analysis of product validation and test plans. *International Journal of Product Development,* **10**(4), pp. 369-392.

ZORRIASSATINE, F., WYKES, C., PARKIN, R. and GINDY, N., 2003. A survey of virtual prototyping techniques for mechanical product development. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture,* **217**(4), pp. 513-530.

Appendix A

Fatigue facts

Fatigue is the progressive structural damage that occurs when metal is subject to cyclic stress. The process starts with a microscopic crack that widens with each stress cycle. As the part continues to undergo repetitive stress, the crack begins to grow rapidly. After a number of stress cycles, the crack grows to critical length at which point, the crack growth becomes unstable and the part fails.

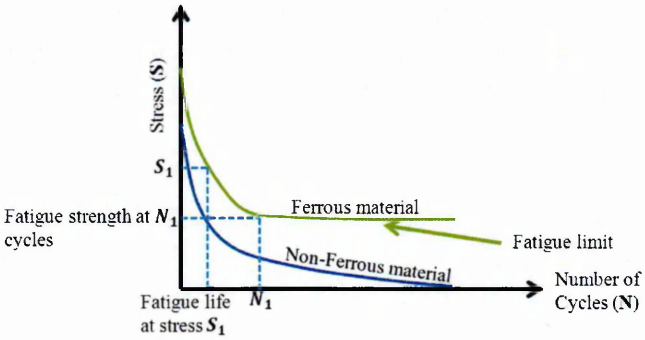


Figure A S-N curve

Engineers use a graph called an "S-N curve" (shows in Figure A) to characterize the fatigue properties of a particular material. The S-N curve plots the magnitude of repetitive stress against the average number of repetitive stress cycles the material can tolerate before it fails. Ferrous metals (like iron and steel) show a "fatigue limit" stress below which they can tolerate an infinite number of repetitive stress cycles without failing, in theory. Non-ferrous metals (like aluminium and copper) have no fatigue limit, and will always fail eventually if subjected to enough stress cycles. Fatigue life is also strongly affected by the magnitude of the repetitive stress cycles. Because a small increase in stress can result in a large decrease in fatigue life. Cylinder heads are usually made of aluminium alloy, so they have a finite fatigue limit. Furthermore, crack growth in a cylinder head can progress rapidly. Fatigues is also affected by corrosions which can initiate fatigue cracks therefore shorten useful fatigue life.